10/506794

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

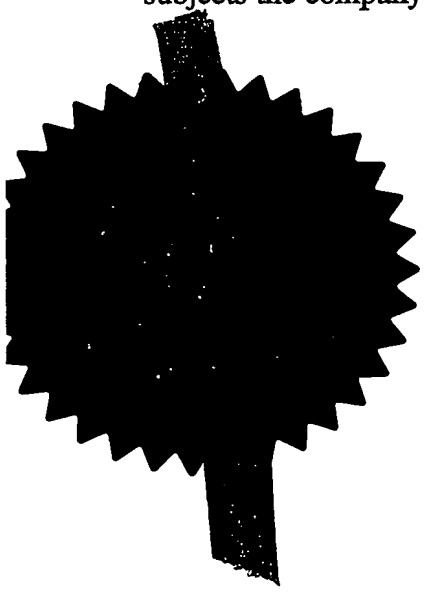
The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

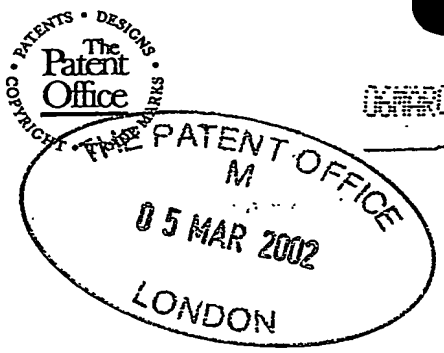


Signed Stephen Hurdle

Dated 19 March 2003

Patents Form 1/77

Patents Act 1977
(Rule 16)



1/77
P01/7700 0.00-0205143.1

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference — P.84806 GCW

2. Patent application number *(The Patent Office will fill in this part)* — **0205143.1** — 5 MAR 2002

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

ARROW THERAPEUTICS LIMITED
Britannia House
7 Trinity Street
London SE1 1DA
United Kingdom

Patents ADP number *(if you know it)* 7708217002

If the applicant is a corporate body, give the country/state of its incorporation — United Kingdom

4. Title of the invention — TRANSPOSON

5. Name of your agent *(if you have one)* — J.A. KEMP & CO.

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

14 South Square
Gray's Inn
London
WC1R 5JJ

Patents ADP number *(if you know it)* 26001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number.

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Yes

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	-
Description	48 ✓
Claim(s)	9 ✓
Abstract	1 ✓
Drawing(s)	6 + 6 ✓

10. If you are also filing any of the following, state how many against each item.

Priority documents	-
Translations of priority documents	-
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	-
Request for preliminary examination and search *(Patents Form 9/77)*	-
Request for substantive examination *(Patents Form 10/77)*	-
Any other documents *(please specify)*	-

11. I/We request the grant of a patent on the basis of this application.

Signature J.A. KEMP & CO. Date 5 March 2002

12. Name and daytime telephone number of person to contact in the United Kingdom

G.C. Woods
020 7405 3292

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

TRANSPOSON

Field of the Invention

This invention relates to a new type of transposon. The transposon may be used in the identification of essential and conditional essential genes, in particular in bacteria.

Background to The Invention

The increase in prevalence of antibiotic-resistant bacteria, for example, has renewed interest in the search for new targets for antibacterial agents. Essential genes and in particular the proteins which they encode may be good substrates for use in screens for antibacterials, antiparasitics, fungicides, pesticides and herbicides. Essential genes and their protein products potentially represent such targets.

Additionally, there is an interest in the identification of conditional essential genes, that is genes which are essential for the survival of an organism in a particular environment. In the case of pathogenic bacteria, for example, conditional essential genes include those which are required for survival in a host. Such genes and the proteins which they encode may also be good targets for use in screens for antibacterials. In particular, bacteria which carry mutations in such genes may be useful in attenuated live vaccines.

Summary of The Invention

Essential genes are those genes which, when missing (eg. because of a chromosomal deletion) or mutated to render them non-functional, result in a lethal phenotype. That is, they are genes without which an organism cannot survive. Conditional essential genes are those genes which, although not absolutely essential for the survival of an organism under all conditions, are essential for survival under various conditional restraints. Examples of particular conditional restraints include survival at elevated temperatures and survival of a pathogen within its host.

A number of transposon-based strategies have been developed for the

identification of essential and conditional essential genes. We have now devised a new transposon. The transposon can be used in a variety of methods for the identification of essential and conditional essential gene in the genome of an organism.

According to the invention there is thus provided a transposon which comprises: (i) an RNA polymerase recognition site; (ii) a homing endonuclease recognition site; and (iii) a bacterial origin of replication.

The invention also provides:

- use of a transposon according to any one of the preceding claims in a method for the identification of an essential or a conditional essential gene;
- a method for identifying an essential gene of an organism, which method comprises:

(i) providing a library of transposon insertion mutants of the said organism, wherein the transposon is a transposon of the invention;

(ii) isolating chromosomal DNA from the library of (i);

(iii) digesting the chromosomal DNA with a restriction endonuclease that is capable of cutting 5' of the RNA polymerase recognition site(s) in the transposon and 3' of the RNA polymerase recognition site(s) in the chromosomal DNA flanking the transposon;

(iv) transcribing RNA from the digested DNA;

(v) hybridising the transcribed RNA with an oligonucleotide array; and

(vi) identifying a probe on the oligonucleotide array which corresponds to an essential gene of the organism;

- a method for identifying an essential gene of an organism, which method comprises:

(i) providing a library of transposon insertion mutants of the said organism, wherein the transposon is a transposon of the invention;

(ii) isolating chromosomal DNA from the library of (i);

(iii) digesting the chromosomal DNA with a homing endonuclease;

(iv) ligating the digested DNA with a biotinylated linker;

(v) recovering the digested DNA using streptavidin-coated particles;

(vi) transcribing RNA from the recovered digested DNA;

(vii) hybridising the transcribed RNA with an oligonucleotide array; and

(viii) identifying a probe on the oligonucleotide array which corresponds to an essential gene of the organism;

- a method for identifying a conditional essential gene of an organism, which method comprises:

(a) providing a first sample of a library of transposon insertion mutants of the said organism (input library);

(b) providing a second sample of the library and subjecting that sample to a conditional restraint;

(c) collecting the mutants that survive the conditional restraint in step (ii) to give a second library (output library);

(d) carrying out a method according to steps (ii) to (iv) of the first method set out above or according to steps (ii) to (vi) of the second method set out above on the input library from step (a) and on the output library from step (c);

(e) hybridising the transcribed RNA derived from the input library and from the output library to the same oligonucleotide array or different separately to copies of the same oligonucleotide array; and

(f) identifying a probe on the oligonucleotide array(s) which corresponds to a conditional essential gene of the organism;

- use of an essential or conditional essential gene identified by a method according to a method of the invention, or a polypeptide encoded by a said gene, in a method for identifying an inhibitor of transcription and/or translation of that gene and/or activity of a polypeptide encoded that gene;

- a method for identifying an inhibitor of transcription and/or translation of an essential or conditional essential gene and/or an inhibitor of activity of a polypeptide encoded by a said gene, which method comprises:

(a) identifying an essential or conditional essential gene by a method according to a method of the invention; and

(b) determining whether a test substance can inhibit transcription and/or translation of a gene identified in step (a) and/or activity of a polypeptide encoded by a said identified gene, thereby to identify a said inhibitor;

- an inhibitor identified by such a method;
- an inhibitor of the invention for use in a method of treatment of a bacterial, fungal or eukaryotic parasite infection, wherein the essential or conditional essential gene used to identify the inhibitor is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene;
- use of such an inhibitor in the manufacture of a medicament for use in the treatment of a bacterial, fungal or eukaryotic parasite infection;
- a pharmaceutical composition comprising such an inhibitor and a pharmaceutically acceptable carrier or diluent;
- a method of treating a host suffering from a bacterial, fungal or eukaryotic parasite infection, which method comprises the step of administering to the host a therapeutically effective amount of such an inhibitor;
- a method for the preparation of a pharmaceutical composition, which method comprises:

(a) identifying an inhibitor of transcription and/or translation of an essential or conditional essential gene of an organism and/or an inhibitor of activity of a polypeptide encoded by a said gene by a method as set out above wherein the essential or conditional essential gene is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene; and

(b) formulating the inhibitor thus identified with a pharmaceutically acceptable carrier or diluent;

- a method for treating a host suffering from a bacterial, fungal or eukaryotic parasite infection, which method comprises:

(a) identifying an inhibitor of transcription and/or translation of an essential or conditional essential gene of an organism and/or an inhibitor of activity of a polypeptide encoded by a said gene by a method as set out above wherein the essential or conditional essential gene is a bacterial, fungal or eukaryotic parasite

essential or conditional essential gene;

(b) formulating the inhibitor thus identified with a pharmaceutically acceptable carrier or diluent; and

(c) administering to the host a therapeutically effective amount of an inhibitor thus formulated;

- an inhibitor of the invention, wherein the essential or conditional essential gene is a plant bacterial, plant fungal or plant pest essential or conditional essential gene;
- use of such an inhibitor as a plant bactericide, fungicide or pesticide;
- an inhibitor of the invention, wherein essential or conditional essential gene is a plant essential or conditional essential gene;
- use of such an inhibitor as a herbicide;
- a bacterium attenuated by a non-reverting mutation in one or more genes identified by a method of the invention;
- a vaccine comprising such a bacterium and a pharmaceutically acceptable carrier or diluent;
- a bacterium as described above for use in a method of vaccinating a human or animal;
- use of such a bacterium for the manufacture of a medicament for vaccinating a human or animal;
- a method for raising an immune response in a mammalian host, which method comprises the step of administering to the host a bacterium as set out above;
- a method for preparing an attenuated bacterium, which method comprises:

(a) identifying a conditional essential gene in a bacterium by a method of the invention; and

(b) introducing a non-reverting mutation into a thus-identified conditional essential gene of the bacterium, thereby to attenuate the bacterium;

- a method for the preparation of a vaccine, which method comprises:

(a) identifying a conditional essential gene in a bacterium by a method of the invention;

(b) introducing a non-reverting mutation into a thus-identified conditional essential gene of the bacterium, thereby to attenuate the bacterium; and

(c) formulating the attenuated bacterium with a pharmaceutically acceptable carrier or diluent.

- a method for raising an immune response in a mammalian host, which method comprises:

(a) identifying a conditional essential gene in a bacterium by a method of the invention;

(b) introducing a non-reverting mutation into a thus-identified conditional essential gene of the bacterium, thereby to attenuate the bacterium;

(c) formulating the attenuated bacterium with a pharmaceutically acceptable carrier or diluent; and

(d) administering to the host the attenuated bacterium thus formulated.

Brief Description of The Figures

Figure 1 shows a diagrammatic representation of the "Gene Kelly" transposon consisting of ME (mosaic end IS sequences), binding sites for the SP6 and T7 RNA polymerases, homing endonuclease sites for I-*Sce*I and PI-*Psp*I, an R6k origin of replication and a kanamycin resistance cassette.

Figure 2 sets out the sequence of the "Gene Kelly" transposon.

Figure 3 shows a diagrammatic representation of the original epicentre EZ:Tn R6k ori Kan transposon with the oligonucleotide sequences overlaid. Black boxes represent matching sequence between PCR oligonucleotides and the epicentre transposon.

Figure 4 sets out a graph showing the position of insertion in the LT2 genome of the 46 sequenced transposon mutants (position in base pairs in increasing number order against number of transposons). From the graph it can be seen that there is a random distribution of insertions throughout the genome.

Figure 5 sets out a diagram of each end of the transposon showing the relative position of the iPCR oligonucleotides with the restriction endonuclease cut sites and

RNA polymerase promoters in transposed chromosomal DNA, which has been digested and religated.

Figure 6 shows a schematic diagram of the ligation capture method of recovering the ends of the transposon.

Description of the sequence listing

SEQ ID NO: 1 sets out the sequence of the "Gene Kelly" transposon.

SEQ ID NO: 2 sets out the sequence of primer 97, which was used in the construction of the "Gene Kelly" transposon.

SEQ ID NO: 3 sets out the sequence of primer 98, which was used in the construction of the "Gene Kelly" transposon.

SEQ ID NO: 4 sets out the sequence of primer 107, which can be used to carry out iPCR in a protocol for generating RNA run-offs.

SEQ ID NO: 5 sets out the sequence of primer 115, which can be used to carry out iPCR in a protocol for generating RNA run-offs.

SEQ ID NO: 6 sets out the sequence of primer 116, which can be used to carry out iPCR in a protocol for generating RNA run-offs.

SEQ ID NO: 7 sets out the sequence of primer 108, which can be used to carry out iPCR in a protocol for generating RNA run-offs.

SEQ ID NO: 8 sets out the sequence of primer 117, which can be used to carry out iPCR in a protocol for generating RNA run-offs.

SEQ ID NO: 9 sets out the sequence of primer 118, which can be used to carry out iPCR in a protocol for generating RNA run-offs.

SEQ ID NO: 10 sets out the sequence of primer 113, which can be used in a protocol for ligation capture recovery of "Gene Kelly" transposon ends.

SEQ ID NO: 11 sets out the sequence of primer 114, which can be used in a protocol for ligation capture recovery of "Gene Kelly" transposon ends.

Detailed Description of The Invention

The invention relates to a new transposon, named the "Gene Kelly" transposon, which is suitable for use in methods for the identification of essential and conditional essential genes.

The transposon of the invention is typically a modified *Tn*5 transposon, although, in principle, any transposon may be modified according to the invention. The transposon of the invention has a combination of features which make it a versatile tool for use in a number of protocols for the identification of essential and conditional essential genes, methods which we refer to as transposon mediated differential hybridisation (TMDH) techniques.

The Gene Kelly transposon comprises one or more RNA polymerase recognition sites (sometimes referred to as RNA polymerase recognition sequences). Typically, the RNA polymerase recognition sequences are located proximal to the end(s) of the transposon, for example adjacent to the mosaic ends (if the transposon has them). The RNA polymerase recognition sequences are typically oriented so that they direct transcription out of the transposon itself. Thus, DNA sequence flanking the integration site of the Gene Kelly transposon may be transcribed.

The transcribed sequences originate from DNA sequences flanking a Gene Kelly transposon. Such DNA sequences are susceptible to insertion and therefore are unlikely to represent essential gene sequences.

Therefore, flanking sequences may be isolated from a library of Gene Kelly insertion mutants and hybridised with an oligonucleotide array which comprises probes corresponding to all of the open reading frames from the organism in question. If the library of Gene Kelly insertion mutants comprises a transposon insertion in all of the non-essential genes of the organism, any probe in the oligonucleotide array to which none of the flanking sequences hybridise is likely to be a good candidate for originating from an essential gene.

Typically, the Gene Kelly transposon comprises two different RNA polymerase recognition sites located proximal to the ends of the transposons both pointing out of the transposon. The use of two different RNA polymerase

recognition sites allows two pools of transcribed RNAs to be isolated from a library of Gene Kelly insertion mutants. One pool will correspond to DNA sequences flanking one side of the transposon insertion site and the other pool will represent sequences flanking the other side of the transposon insertion site. The generation of these two pools may help to minimise the risk of essential gene being incorrectly assigned as non-essential genes. This is explained in more detail below.

The Gene Kelly transposon also comprises homing endonuclease recognition sites. Homing endonucleases are rare cutters, especially in bacterial DNA. Incorporation of recognition sites for such endonucleases into a genome effectively permits the introduction of these sites into the genome being studied. Transposed DNA may be digested with the appropriate homing endonuclease and the resulting ends (if none are present in the bacterial genome) should therefore all originate from the Gene Kelly transposon. Those ends may be rescued using a ligation-capture technique described below allowing the rapid and selective purification of regions of the genomic DNA of the organism being studied which originate from a site of transposon insertion. Such regions can be used in hybridisation experiments with oligonucleotide arrays as outlined above.

The use of two homing endonuclease recognition sites in combination with two RNA polymerase recognition sites allows sequences flanking the two sides of the inserted transposons to be isolated separately. The generation of these two pools may help to minimise the risk of essential gene being incorrectly assigned as non-essential genes. This is explained in more detail below.

The third critical feature in the Gene Kelly transposon is the incorporation of a bacterial origin of replication. This allows plasmid-rescue to be carried out of the complete transposon (plus the flanking regions of chromosomal DNA around the insertion site). The protocol requires chromosomal digestion of genomic DNA isolated from a library of Gene Kelly insertion mutants (with a restriction endonuclease that does not cut in the tranposon sequence or at least in the bacterial origin of replication in the tranposon sequence), religation and then transformation into a strain of bacteria in which the origin of replication will function.

The Gene Kelly transposon thus comprises three critical features: (i) an RNA polymerase recognition site; (ii) a homing polymerase recognition site; and (iii) a bacterial origin of replication. The full sequence of the Gene Kelly transposon is set out in Figure 2 and SEQ ID NO: 1.

Transposons, sometimes called transposable elements, are mobile polynucleotides. The term transposon is well known to those skilled in the art and includes classes of transposons that can be distinguished on the basis of sequence organisation, for example short inverted repeats at each end; directly repeated long terminal repeats (LTRs) at the ends; and polyA at 3' ends of RNA transcripts with 5' ends often truncated. Some types of virus also integrate into the host genome, for example retroviruses, and may therefore be used to generate libraries of insertion mutants. However, transposons are typically preferred to viruses because issues of safety related to pathogenicity may be avoided.

Any suitable transposons may be modified to generate a Gene Kelly transposon, as long as it comprises the three features set out above. Suitable transposons for use in bacteria which can be modified to generate a Gene Kelly transposon include *Tn*3, γδ, *Tn*10, *Tn*5, *Tn*phoA, *Tn*903, *Tn*917, Bacteriophage Mu and related viruses. Any of the above mentioned transposons may be modified to generate a Gene Kelly transposon.

Preferred Gene Kelly transposons are those which carry antibiotic resistance genes (which may be useful in identifying mutants which carry a transposon) conferring resistance to, for example, kanamycin, streptomycin and bleomycin. A particular preferred transposon for use in the invention is a modified *Tn*5 transposon.

Suitable transposons for use in fungi which can be modified to generate a Gene Kelly transposon include the *Ty*1 element of *Saccharomyces cerevisiae*, the filamentous fungi elements (the filamentous fungi include agriculturally important plant pathogens such as *Erysiphe graminis*, *Magnaporthe grisea*) such as *Fot*1/*Pogo*-like and *Tc*1/*Mariner*-like elements (see Kempen and Kuck, 1998, Bioessays 20, 652-659 for a review of such elements).

Suitable transposons for use in plants which can be modified to generate a

Gene Kelly transposon include *Ac/Ds*, *Tam*3 and other *Tam* elements, *cin*4 and *spm*.

Suitable transposons for use in animals which can be modified to generate a Gene Kelly transposon include *P* and *hobo* which may be used in *Drosophila* and *Tc*1 which can be used in *Caenorhabditis elegans*.

The Gene Kelly transposon a recognition site for an RNA polymerase. Preferred recognition sites are those for which the corresponding RNA polymerase is highly selective for initiation. Other preferred recognition sites are those for which the corresponding RNA polymerase does not initiate transcription from sequences of the organism being studied. Preferred examples of RNA recognition sites are those recognised by bacteriophage RNA polymerases, for example the recognition site for T7 RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase, in particular the T7 RNA polymerase recognition site. The recognition sites for these particular RNA polymerases are well known to those skilled in the art.

The RNA polymerase recognition site may appear anywhere within a transposon for use in the invention. However, typically the RNA polymerase recognition site will be located proximal to one end of the transposon, i.e. proximal to one IS/ME. Typically, the 3' end of the RNA polymerase recognition site will be situated from one to 30, for example from five to twenty base pairs away from the 5' end of one of the IS/ME.

Preferred Gene Kelly transposons may comprise two RNA polymerase recognition sites, which generally will be different RNA polymerase recognition sites. For example, a Gene Kelly transposon may comprise a T7 RNA polymerase recognition site and an SP6 RNA polymerase recognition site. A Gene Kelly transposon may also comprise more than two RNA polymerase recognition sites, for example three or four RNA recognition sites.

A Gene Kelly transposon also comprises a homing endonuclease recognition site. Homing endonucleases are also known as intron or intein encoded endonuclease. They are encoded by genes with mobile, self-splicing introns or inteins (protein introns). Any homing nuclease recognition site may be used, for example I-*Sce*I, PI-*Psp*I or I-*Ppo*I.

The homing endonuclease recognition site may be located anywhere in the transposon, but typically it is situated 5' to an RNA polymerase recognition site, i.e. further into the transposon than the RNA polymerase recognition site. Preferred Gene Kelly transposons may comprise two homing endonuclease recognition sites which generally will be different homing endonuclease recognition sites, for example a I-*Sce*I and a PI-*Psp*I recognition site. A Gene Kelly transposon may, however, comprise more than two homing endonuclease recognition sites, for example three or four homing endonuclease recognition sites.

Typically, the Gene Kelly transposon will comprise the same number of RNA polymerase recognition sites as homing endonuclease recognition sites, ideally two of each. Each homing endonuclease recognition site will typically be located 5' to an RNA polymerase recognition site, for example such that the 3' end of the homing endonuclease recognition site is from one to 30, for example, from five to twenty base pairs away from the 5' end of one of an RNA polymerase recognition site.

The Gene Kelly transposon also comprises a bacterial origin of replication. A preferred example of a suitable bacterial origin of replication is the R6k origin of replication. The R6k origin of replication is capable of functioning in a *pir*$^+$ strain of bacteria.

The Gene Kelly transposon may be used in methods for the identification of essential or conditional essential genes, i.e. use of a transposon of the invention in a method for the identification of an essential gene or a conditional essential gene is provided according to the invention. The methods described below for the identification of essential or conditional essential genes using the Gene Kelly transposon can conveniently be referred to as transposon mediated differential hybridisation (TMDH).

In methods for the identification of essential gene, typically the first step is the provision of a library of transposon insertion mutants. Libraries of Gene Kelly transposon insertion mutants may be generated according to any method known to those skilled in the art. For example, libraries of bacterial transposon insertion mutants can be constructed using either plasmid or bacteriophage vectors containing

the transposon and a selectable marker. Bacteriophage λ, eg. λ*Tnpho*A can be used to infect a suitable recipient bacterial strain, for example *E. coli* XAC. This *E. coli* strain has a suppressor mutation which prevents the bacteriophage from replicating and subsequently lysing and also contains an antibiotic resistance gene to allow selection of colonies containing transposed chromosomal DNA. The vector contains mutation(s) preventing integration of the λ chromosome into the bacterial host chromosome and thus the growth of false positive colonies without a mutated *E. coli* gene is prevented. Cultures of the recipient strain are grown in enriched medium (eg. Luria Broth) and cells in mid log phase of growth are infected with the λ transposon vector for 1 hour at 37°C. Aliquots of the infected cells are plated out on L-agar supplemented with the appropriate selective antibiotic and grown overnight at 37°C. These colonies constitute a transposon library and can be further analysed by the TMDH procedure described in this application.

Alternatively, transposasome complexes comprising the transposon in a complex with a transposase may be generated and electroporated into a suitable electrocompetent host. Suitable techniques for preparing transposasomes and for electroporating transposasomes into host cells are well known to those skilled in the art.

Growth of such libraries results in the generation of potentially thousands of insertion mutants all of which mutants carry insertion that are, of necessity, in genes that (when mutated) do not result in the death of the cell ie. are non-essential genes.

Each mutant in a suitable transposon insertion library may carry one transposon insertion. However, a mutant may carry more than one transposon insertion, for example two, three, four, five, ten or twenty transposon insertions. A transposon insertion mutant library suitable for use in the invention will comprise at least one transposon insertion mutant for at least 60%, at least 70%, typically at least 80%, preferably at least 90%, more preferably at least 95%, even more preferably at least 99%, or most preferably substantially all of the non-essential genes in the organism being studied. Preferably the library will be a saturating library, i.e. the library comprises a transposon insertion mutant for substantially all genes of the

organism that when mutated give rise to a viable organisms for example bacterium. A transposon insertion could be in an open reading frame of a gene or in a regulatory sequence of gene.

Any non-essential gene in the transposon insertion library may be represented by more than one insertion mutant, for example two, three, four, five or up to ten insertion mutants, each carrying transposon insertions at the same or different sites in the non-essential gene or carrying insertions at the same site in different orientations. Preferred libraries will have, on average, more than one different transposon insertion mutant for each non-essential gene represented in the library, for example at least two on average, at least four on average, at least 5 on average or at least 10 on average different transposon insertion mutants for each essential gene represented in the library.

Some regions of a particular genome may be inaccessible to insertion by a particular transposon, for example because of a particular secondary or tertiary structure which is inaccessible to a particular transposon. Thus it may be advantageous to combine two transposon libraries, thereby increasing the probability of obtaining transposon insertions in a greater number of genes. For example, in the case of bacterial libraries, *Tn*5 and *Tn*10 (both modified by the inclusion of an RNA polymerase recognition sequence) libraries for example, could be combined.

RNA target sequences are generated from the DNA (host organism) sequences that flank the transposons, i.e. non-essential gene sequences.

Generally, flanking sequence will be isolated from at least 60%, for example at least 70%, at least 80%, preferably at least 90%, more preferably at least 95% and most preferably at least 99% of the transposon insertion mutants in a particular library of mutants.

In the method of the invention chromosomal DNA is prepared from the library of transposon insertion mutants. Techniques for the isolation of chromosomal DNA, alternatively referred to as genomic DNA, are well known to those skilled in the art. The Gene Kelly transposon allows a number of different techniques to be used for the generation of RNA target sequences from the isolated genomic DNA.

In one version of TMDH, the chromosomal DNA thus prepared is then digested with a restriction endonuclease. The restriction endonuclease is one which is capable of cutting at a recognition site which is located in the transposon at a position 5' to the RNA polymerase recognition site and in the chromosomal DNA flanking the transposon 3' to the RNA recognition site (which is in the transposon). The exact restriction enzyme to be used will depend on the sequence of the transposon. However, typically an restriction endonuclease is used which has recognition sites that appear frequently within the genome of the organism being studied. Thus, a series of DNA fragments is generated, some of which comprise an RNA polymerase recognition site fused to a portion of flanking sequence, i.e. non-essential gene sequence.

Generally, suitable restriction endonucleases will have six base pair, five base pair or preferably four base pair recognition sequences. Suitable examples of four base pair cutters are set out in Table 1 below:

Table 1. Examples of 4bp recognition type II restriction endonucleases suitable for use in TMDH

Enzyme	Recognition Site	Enzyme	Recognition Site
*Aci*I	C^CGC GGC_G	*Mse*I	T^TAA AAT_T
*Alu*I	AG^CT TC_GA	*Msp*I	C^CGG GGC_C
*Bfa*I	C^TAG GAT_C	*Nla*III	^CATG GTAC_
*Bstu*I	CG^CG GC_GC	*Rsa*I	GT^AC CA_TG
*Dpn*I	^GATC CTAG_	*Sau*3a	^GATC CTAG_
*Hae*III	GG^CC CC_GG	*Taq*I	T^CGA AGC_T

*Hinp*I	G$^{\downarrow}$CGC CGC$_{\uparrow}$G	*Tsp*509	$^{\downarrow}$AATT TTAA$_{\uparrow}$
*Hha*I	GCG$^{\downarrow}$C C$_{\uparrow}$GCG		

The resulting fragments comprise an RNA polymerase site adjacent to non-essential gene sequence. The fragments may optionally be size selected. If size selection is carried out, fragments with a size of from about 100 bp to about 2000bp or preferably of from about 200 bp to about 600 bp may be isolated, for example from a gel, and purified. The smaller the fragments isolated, the smaller the chance of the RNA target sequences including sequences from genes which lie next to genes which have been interrupted by transposons. If such adjacent sequences were from essential genes, there is the possibility that essential gene sequences could be identified as non-essential gene sequences. Thus, size fractionation may reduce the amount of false non-essential gene sequences.

The transposon:flanking sequence fragments are then used to generate the RNA target sequences. Following digestion, the transposon:flanking sequence fragments (the "target") can be amplified typically by PCR and preferably by iPCR (inverse PCR). *In vitro* transcription from the RNA polymerase recognition sequence can then be carried out. Linear amplification can alternatively be achieved by *in vitro* transcription without the prior amplification step. Techniques for carrying out *in vitro* transcription are will known to those skilled in the art and any suitable technique may be used. In essence, an RNA polymerase and ribonucleotides are used. Preferably, one or more of the ribonucleotides is labelled so that the resulting RNA target sequences are labelled. Suitable labels include radioactive labels, for example ^{32}P, ^{33}P or ^{35}S, or non-radioactive labels, for example an enzyme, a fluorescent label or biotin.

Fluorescent labels are preferred, for example a water-soluble fluorescent dye such as Cy3™ or Cy5™ or a fluorescein-tagged compound such as FluorX™ (the NHS ester of carboxyfluorescein with an extended linker arm), fluorescein isothiocyanate (FITC) or 5-([4,6-Dichlorotriazin-2-yl]amino)fluorescein (DTAF). Generally, it will only be necessary to have one of the four ribonucleotides labelled.

The techniques described above allow the isolation of sequences flanking the transposons in a library of transposon insertion mutants. Thus, a pool of flanking sequences is generated collectively referred to as the RNA target sequences. Although fragments in the pool are generated from only one side of the transposons, a transposon is capable of inserting at any particular locus (that can be disrupted) in either orientation. Thus, in particular in a saturating transposon insertion library, many loci will be represented by mutants carrying insertions in both orientations. Therefore, the RNA target sequences generated according to the TMDH method of the invention will, for many loci, comprise flanking sequence in both orientations.

In a modification of this technique, the fragments comprising the RNA polymerase site fused to non-essential gene sequence may be self-ligated. Suitable techniques for carrying out self-ligation are well known to those skilled in the art. Any suitable ligase may be used, for example T4 DNA ligase. Ligation reactions may be carried out for from 6 to 24 hours, for example from 12 to 16 hours at a temperature of from 10°C to 20°C, for example at about 16°C.

Self-ligated molecules are then amplified using iPCR. Techniques for carrying out iPCR are well known to those skilled in the art and may be carried out according to any suitable technique. Typically, iPCR is carried out using two oligonucleotides which bind divergently at a location 5' to the RNA polymerase recognition site and 3' to the restriction endonuclease recognition site in the transposon, i.e. the two oligonucleotide recognition sites are located on the transposon between the restriction endonuclease recognition site and the RNA polymerase recognition site.

When using iPCR techniques, there is the possibility that, a "stuffer" fragment may ligate into the self-ligation reaction, which will be amplified along with the transposon-disrupted sequence. If this material were to be used in subsequent generation of the RNA target sequences, the stuffer sequence could create non-specific background signal as it would also be hybridized to the high density array. In order to remove this stuffer fragment, the sequences amplified in iPCR can be redigested with whichever enzyme was used to isolate the transposon-flanking

sequence fragments in the first place. This results in the release of the stuffer fragments which can be removed from the transposon:flanking sequence fragments. Removal of the stuffer fragments can be facilitated if biotinylated primers are used in iPCR. The transposon:flanking sequence fragments can then be removed from the stuffer fragments using a magnetic-bead-streptavidin conjugate.

In a further modification of TMDH method, the chromosomal DNA isolated from a transposon insertion library may be divided into a number of aliquots. Those aliquots may then each be separately digested with a different restriction endonuclease which is capable of cutting at a recognition site which is located in the transposon at a position 5' to the RNA polymerase recognition site and in the chromosomal DNA flanking the transposon 3' to the RNA polymerase recognition site (which is in the transposon). The chromosomal DNA may be separated into, for example two, three, four, five or ten aliquots which are each separately digested with a different restriction endonuclease. Preferred restriction enzymes are as set out in Table 1 above.

Thus, for example, two or three aliquots of the chromosomal DNA may be separately digested with different suitable restriction endonucleases, for example two or three of *Hae*III, *Hha*I, H*pych4*IV and *Rsa*I.

If the TMDH protocol is used in this modified format, the different aliquots may be repooled after digestion and treated together in the subsequent steps of TMDH. Alternatively, the digested aliquots may be treated separately in the subsequent steps. If this TMDH format is adopted, a number of pools of RNA target sequences result. Each pool of RNA target sequences may be labelled with a different, for example fluorescent, label.

In an alternative protocol for the preparation of RNA target sequences, the genomic DNA isolated from a library of Gene Kelly transposons is digested with a homing endonuclease. Homing endonuclease recognition sites are rare and therefore any "ends" generated by digestion with a homing endonuclease should originate from the transposon. Those ends may be rescued by annealing a biotinylated linker to them and then isolating the biotinylated transposon:flanking sequence fragments with

streptavidin-coated particles, for example streptavidin-coated magnetic beads. Alternatively the ends may be rescued by labelling with digoxygenin and isolating the labelled fragments with an antibody raised against digoxygenin. The transposon:flanking sequence fragments are then used to generate the RNA target sequences by carrying out *in vitro* transcription from the RNA polymerase recognition site. *In vitro* transcription can be carried out as described above.

The sequences which comprise the RNA target sequences (whichever of the methods described above is used to generate them) may be used for hybridisation with olignucleotide arrays. Oligonucleotide arrays used in the TMDH protocol of the invention are preferably high density arrays.

Oligonucleotide arrays suitable for use in the invention may comprise sequences from one or more loci of a genome. Preferably suitable oligonucleotide arrays will represent at least 80% of all open reading frames (ORFs), more preferably at least 90% of all ORFs, for example 95% of all ORFs, even more preferably 99% of all ORFs or substantially all ORFs of the genome represented on the oligonucleotide array.

By high density oligonucleotide array is meant an array in which there are a high number of probes covering the locus, loci or genome represented by the array. For example, in a high density oligonucleotide array there may be a probe, for example, for every 30 to 500 base pairs of the locus, loci or genome represented by the array. Preferably there will be a probe, for every 60 to 250 base pairs of locus, loci or genome represented in the array, for example about every 100 base pairs. Probes may overlap, for example by 1, 2, 3, 4, 5, up to 10, up to 20, up to 30, up to 40 or up to 50 bases.

The olignucleotide probes on the array are, for example, from about 8 or 9 to about 150 nucleotides in length, preferably from about 30 or 50 to about 100 nucleotides in length or more preferably about 60 nucleotides in length.

The oligonucleotide probes used in the array will typically be designed on the basis of the wild type sequence of the organism being studied. The oligonucleotide probes may be designed so that each probe has minimal or substantially no cross-

hybridisation with other sequences in the genome from which the probes originate. The BLAST program can be used to design suitable probes (Altschul *et al.*, J. Mol. Biol. 215, 403-410).

Methods for making oligonucleotide arrays are well known to those skilled in the art.

Probes which show no hybridisation or substantially no hybridisation (there may be a low level of background non-specific hybridisation) with the RNA target sequences are unlikely to have been disrupted by a transposon insertion event and consequently are strong candidates for sequences corresponding to essential genes.

However, it is theoretically possible for oligonucleotide probes within the 5' or 3'-termini of essential genes to show a hybridisation signal with the TMDH protocol. For example, if a transposon insertion occurs in a non-essential gene adjacent to an essential gene, RNA target sequences may be generated from this transposon corresponding to both non-essential and essential gene sequences as a result of restriction sites lying within the essential gene. The resulting labelled target will not only comprise DNA corresponding to the non-essential gene (that has been disrupted), but will also extend into the adjacent essential gene up to the restriction site. The result of hybridising this labelled target to the oligonucleotide array will be appear as "bleed through" of signal to probes on either the 5' or 3' end of the essential gene, up to the restriction site used for the TMDH protocol.

To address this potential source of mis-assignment of essential genes, the restriction endonuclease digestion TMDH protocol described above may be carried with more than one aliquot of the isolated genomic DNA, for example two or three, whereby each aliquot is digested with a different restriction endonuclease (which have different recognition motifs). The more aliquots digested with more restriction sites that are used to generate target sequences, the more statistically unlikely it is that all of them will result in labelled RNA target sequences that "bleed through" into essential genes. The pools of RNA target sequences derived from the different digestions can be hybridised to the same or, preferably, a different oligonucleotide array if they were generated using different labels, or alternatively may be hybridised

to copies of the same or, preferably, a different array. The analysis of the resulting multiple array hybridisation patterns will remove any ambiguity on the site of transposon invention.

A different modification to the TMDH protocol, applicable to both the standard restriction endonuclease and the homing endonuclease approaches, to minimize the mis-assignment of essential genes as non-essential gene is to isolate two pools of RNA target sequences which each originate from different sides of the transposons. This approach is illustrated in Figures 5 and 6. Figures 5 and 6 show Gene Kelly transposons which comprise two RNA polymerase recognition sequences (although in Figure 6 the two ends are shown superimposed).

In figure 5 one pool of RNA target sequences (from one side of the transposons) is generated by carrying out *in vitro* transcription using T7 polymerase and a second pool of RNA target sequences (from the other side of the transposon) is generated by carrying out *in vitro* transcription using T3 or SP6 polymerase.

Similarly, in figure 6 one pool of RNA target sequences (from one side of the transposons) is generated by carrying out *in vitro* transcription using T7 polymerase and a second pool of RNA target sequences (from the other side of the transposon) is generated by carrying out *in vitro* transcription using T3 or SP6 polymerase. The ends lying 5' to the RNA polymerase recognition site (which are rescued using streptavidin coated particles) are generated via digestion with homing endonucleases. The homing endonuclease recognition site associated with each RNA polymerase recognition site may be the same or different homing endonuclease recognition sites.

The two pools of RNA target sequences can be hybridised to the same oligonucleotide array if they were generated using different labels, or alternatively each may be hybridised to a separate copy of the same array. A probe in the oligonucleotide array which shows no hybridisation to either pool is likely to correspond to an essential gene.

Where essential gene sequence is isolated in one of the pools because it lies close to non-essential gene sequence flanking a transposon, hybridisation to a probe on the oligonucleotide array will be observed even though that probe corresponds to

an essential gene. However, that probe will show no hybridisation with the other pool of RNA target sequences which comprise the flanking sequence from the other side of the transposon. Thus, in this type of TMDH, a probe in the oligonucleotide array to which at least one of the pools of RNA target sequences does not hybridise is likely to correspond to a gene that has not been disrupted by a transposon and may therefore be an essential gene.

The TMDH methods described above may also be used for the identification of conditional essential genes. Conditional essential genes are those which are not absolutely essential for bacterial survival, but are essential for survival in particular environments eg. survival, i.e. growth/proliferation, in a host (in the case of a pathogenic bacterium) or survival at elevated temperatures. Such environments are known as conditional restraints.

In order to isolate conditional essential genes, a library of transposon mutants is generated under control conditions (eg. growth at 37°C in complete media). The library of mutants is then subjected to some conditional restraint. For example, the library of mutants can be inoculated in a suitable host, if it is a pathogen. Alternatively, the library of mutants can be grown at an elevated temperature. After the library of mutants has been subjected to the conditional restraint it can be recovered.

The library of mutants that have been exposed to the conditional restraint will lack mutants which carry transposons in those genes essential for growth under the conditional environment, for example growth/proliferation in a host organism.

The control and conditional restraint libraries can then be subjected to the TMDH protocols described above using the Gene Kelly transposon. The two resulting RNA target sequence libraries may then be hybridised separately to high density oligonucleotide arrays. Alternatively they can be hybridised to the same array, if the control and conditional restraint libraries are differentially labelled for example.

Comparison of the results given with the control and the conditional restraint libraries will allow the identification of genes which permit survival in the

conditional restraint. Genes identified as essential for survival in the conditional restraint library, but not identified as essential for survival under control conditions should represent genes that are essential for survival under the conditional restraint. In particular, probes which show hybridisation with RNA target sequences from the input library but which show no hybridisation or substantially no hybridisation (there may be a low level of background non-specific hybridisation) with RNA target sequences from the output library are strong candidates for sequences corresponding to conditional essential genes. The same "bleed through" considerations apply as set out above and the modified TMDH protocols for overcoming such "bleed through" may need to be used.

In the case of the analysis of conditional mutations in a pathogen, a library of *Salmonella typhimurium* transposon mutants, for example, can be used to infect a mouse. Following infection, bacteria target to livers and spleens and the course of infection can be conveniently followed by performing viable bacterial counts on those organs. The bacteria recovered from the livers and spleens can be grown on suitable plates. In the case of the conditional restraint at elevated temperature, a transposon-tagged library can be grown at 42°C.

Other conditional restraints include growth of antibiotic resistant bacteria in the presence of antibiotics. This may reveal genes which are essential for antibiotic resistance. Such genes would be targets for drugs with the ability to lower bacterial resistance to particular antibiotics. Organisms could be grown in the presence of carcinogens, UV or other agents that cause oxidative stress and thus genes that confer resistance to growth under those conditions may be identified.

Potential essential gene sequences and conditional essential gene sequences identified by a TMDH strategy using the Gene Kelly transposon may be verified using a method based on allelic exchange. This technique is particularly suitable for analysis of bacterial genes. PCR primers can be used to generate left- and right-arm sequences corresponding to the target gene sequence and ligated with a kanamycin-resistance encoding gene cassette. The resulting cassette can be introduced into a suicide vector, for example a plasmid-based vector, which is unable

to replicate in a host bacterium.

In the case of a candidate essential gene, the resulting construct can be introduced into the bacterial strain from which the candidate gene originates. If the target gene is essential, it should be impossible to isolate allelic-exchange mutants that have a disrupted version of the target gene. In the case of a candidate conditional essential gene, the essential gene can be introduced into the bacterial strain from which the candidate gene originates. Allelic-exchange mutants can be isolated and subjected to growth under the conditional restraint. If the candidate gene is a conditional essential gene, it should not be possible for the allelic-exchange mutants to survive under the conditional restraint.

Similar experiments may be performed for other organisms.

The use of bioinformatics may allow the rapid isolation of further essential and conditional essential genes. A gene identified by TMDH using the Gene Kelly transposon may be used to search databases containing sequence information from other species in order to identify orthologous genes from those species. Genes so identified can be tested for being essential or conditionally essential using the genetic techniques described above. For example, an *E. coli* gene is identified as essential using a method as described above. This may allow the identification of a putative orthologue from Salmonella. That Salmonella gene may be tested by allelic exchange and the construction of conditional mutants in Salmonella as described above. Further orthologues may be identified in more distantly related organisms, for example from Plasmodium species.

Suitable bioinformatics programs are well known to those skilled in the art. For example, the Basic Local Alignment Search Tool (BLAST) program (Altschul *et al.*, 1990, J. Mol. Biol. 215, 403-410. and Altschul *et al.*, 1997, Nucl. Acids Res. 25, 3389-3402.) may be used. Suitable databases for searching are for example, EMBL, GENBANK, TIGR, EBI, SWISS-PROT and trEMBL.

Organisms that may be used in the invention are those for which it is possible to carry out transposon mutagenesis and thus, those that can give rise to a library of transposon mutants. Clearly, if the genome is bigger, more mutants will have to be

produced in order to give a better chance of achieving saturation mutagenesis. Suitable organisms include prokaryotic and eukaryotic organisms. Suitable prokaryotes include bacteria. Preferred bacteria are those which are animal or human or plant pathogens.

The bacteria used may be Gram-negative or Gram-positive. The bacteria may be for example, from the genera *Escherichia, Salmonella, Vibrio, Haemophilus, Neisseria, Yersinia, Bordetella, Brucella, Shigella, Klebsiella, Enterobacter, Serracia, Proteus, Vibrio, Aeromonas, Pseudomonas, Acinetobacter, Moraxella, Flavobacterium, Actinobacillus, Staphylococcus, Streptococcus, Mycobacterium, Listeria, Clostridium, Pasteurella, Helicobacter, Campylobacter, Lawsonia, Mycoplasma, Bacillus, Agrobacterium, Rhizobium, Erwinia* or *Xanthomonas*. Examples of some of the above mentioned genera are *Escherichia coli* - a cause of diarrhoea in humans; *Salmonella typhimurium* - the cause of salmonellosis in several animal species; *Salmonella typhi* - the cause of human typhoid fever; *Salmonella enteritidis* - a cause of food poisoning in humans; *Salmonella choleraesuis* - a cause of salmonellosis in pigs; *Salmonella dublin* - a cause of both a systemic and diarrhoeal disease in cattle, especially of new-born calves; *Haemophilus influenzae* - a cause of meningitis; *Neisseria gonorrhoeae* - a cause of gonorrhoea; *Yersinia enterocolitica* - the cause of a spectrum of diseases in humans ranging from gastroenteritis to fatal septicemic disease; *Bordetella pertussis* - the cause of whooping cough; *Brucella abortus* - a cause of abortion and infertility in cattle and a condition known as undulant fever in humans; *Vibrio cholerae* - a cause of cholera; *Clostridium tetani* - a cause of tetanus; *Bacillus anthracis* - a cause of anthrax. Suitable eukaryotes include fungi, plants and animals. Preferred eukaryotes include animal or human parasites and plant pests.

Suitable fungi include the animal pathogens including *Candida albicans* - a cause of thrush, *Trichophyton* spp. - a cause of ringworm in children, athlete's foot in adults. Other suitable fungi include the plant pathogens *Phytophthora infestans, Plasmopara viticola, Peronospora* spp., *Saprolegnia* spp., *Erysiphe* spp., *Ceratocystis ulmi, Monilinia fructigena, Venturia inequalis, Claviceps purpurea,*

Diplocarpon rosae, *Puccinia graminis*, *Ustilago avenae*.

Suitable animal parasites include *Plasmodium* spp., *Trypanasoma* spp., *Giarda* spp., *Trichomonas* spp. and *Schistosoma* spp. Other animal parasites include the various platyhelminth, nematode and annelid parasites.

Suitable plant pests include insects, nematodes and molluscs such as slugs and snails.

Suitable plants include monocotyledons and dicotyledons.

Preferred organisms are those for which the entire genome is known and for which it may be possible to construct a high density oligonucleotide array covering the entire genome or all of the open reading frames.

Essential and conditional essential genes of bacteria and the polypeptides which they encode may represent targets for antibacterial substances. Similarly essential and conditional essential genes of fungi and eukaryotic parasites, pests and plants and the proteins which they encode may represent targets for fungicides, antiparasitics, pesticides and herbicides respectively. Fungicides may have both animal and plant applications.

Furthermore, if a particular gene is essential or conditionally essential for a number of different bacteria, fungi, parasites, pests or plants, that gene and the polypeptide it encodes may represent a target for substances with a broad-spectrum of activity.

An essential or conditional essential gene identified by one of the methods described above using the Gene Kelly transposon and the polypeptide which it encodes may be used in a method for identifying an inhibitor of transcription and/or translation of the gene and/or activity of the polypeptide encoded by the gene. Such a method may comprise identifying an essential or conditional essential gene using a method of the invention and then determining whether a test substance inhibits the transcription and/or translation of a gene thus identified or inhibits the activity of a polypeptide encoded by such a gene.

Such a substance may be referred to as an inhibitor of an essential or conditional essential gene. Thus, an inhibitor of an essential or conditional essential

gene is a substance which inhibits expression and/or translation of that essential gene and/or activity of the polypeptide encoded by that essential or conditional essential gene.

Any suitable assay may be carried out to determine whether a test substance is an inhibitor of an essential or conditional essential gene. For example, the promoter of an essential or conditional essential gene may be linked to a coding sequence for a reporter polypeptide. Such a construct may be contacted with a test substance under conditions in which, in the absence of the test substance expression of the reporter polypeptide would occur. This would allow the effect of the test substance on expression of the essential or conditional essential gene to be determined.

Substances which inhibit translation of an essential or conditional essential gene may be isolated, for example, by contacting the mRNA of the essential or conditional essential gene with a test substance under conditions that would permit translation of the mRNA in the absence of the test substance. This would allow the effect of the test substance on translation of the essential or conditional essential gene to be determined.

Substances which inhibit activity of a polypeptide encoded by the essential gene may be isolated, for example, by contacting the polypeptide with a substrate for the polypeptide and a test substance under conditions that would permit activity of the polypeptide in the absence of the test substance. This would allow the effect of the test substance on activity of the polypeptide encoded by the essential or conditional essential gene to be determined.

Suitable control experiments can be carried out. For example, a putative inhibitor should be tested for its activity against other promoters, mRNAs or polypeptides to discount the possibility that it is a general inhibitor of gene transcription, translation or polypeptide activity.

Suitable test products which can be tested in the above assays include combinatorial libraries, defined chemical entities, peptide and peptide mimetics, oligonucleotides and natural product libraries, such as display (e.g. phage display

libraries) and antibody products. Antibody products include monoclonal and polyclonal antibodies, single chain antibodies, chimaeric antibodies and CDR-grafted antibodies.

Typically, organic molecules will be screened, preferably small organic molecules which have a molecular weight of from 50 to 2500 daltons. Candidate products can be biomolecules including, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs.

Test substances may be used in an initial screen of, for example, for example from 10 to 100 substances per reaction, and the substances of these batches which show inhibition or stimulation tested individually. Test substances may be used at a concentration of from 1nM to 1000μM, preferably from 1μM to 100μM, more preferably from 1μM to 10μM. Suitable test substances for inhibitors of essential or conditional essential genes include combinatorial libraries, defined chemical entities, peptides and peptide mimetics, oligonucleotides and natural product libraries.

The test substances may be used in an initial screen of, for example, ten substances per reaction, and the substances of batches which show inhibition tested individually.

An inhibitor of an essential or conditional essential gene is one which inhibits expression and/or translation of that essential gene and/or activity of the polypeptide encoded by that essential or conditional gene. Preferred inhibitors of the invention are those which inhibit essential gene expression and/or translation and/or activity by at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95% or at least 99% at a concentration of the inhibitor of 1 $ngml^{-1}$, 10 $ngml^{-1}$, 100 $ngml^{-1}$, 500 $ngml^{-1}$, 1 μgml^{-1}, 10 μgml^{-1}, 100 μgml^{-1}, 500 μgml^{-1}, 1 $mgml^{-1}$, 10 $mgml^{-1}$, 100$mgml^{-1}$. The percentage inhibition represents the percentage decrease in expression and/or translation and/or activity in

represents the percentage decrease in expression and/or translation and/or activity in a comparison of assays in the presence and absence of the test substance. Any combination of the above mentioned degrees of percentage inhibition and concentration of inhibitor may be used to define an inhibitor of the invention, with greater inhibition at lower concentrations being preferred.

Test substances which show activity in assays such as those described above can be tested in *in vivo* systems, such as an animal model of infection for antibacterial activity or a plant model for herbicidal activity. Thus, candidate inhibitors could be tested for their ability to attenuate bacterial infections in mice in the case of an antibacterial or for their ability to inhibit growth of plants in the case of a herbicide.

Inhibitors of bacterial, fungal or eukaryotic parasite essential or conditional essential genes may be used in a method of treatment of the human or animal body by therapy. In particular such substances may be used in a method of treatment of a bacterial, fungal or eukaryotic parasite infection. Such substances may also be used for the manufacture of a medicament for use in the treatment of a bacterial, fungal or eukaryotic parasite infections The condition of a patient suffering from such an infection can be improved by administration of an inhibitor. A therapeutically effective amount of an inhibitor may be given to a human patient in need thereof. Inhibitors of bacterial, fungal or eukaryotic parasite essential or conditional essential genes may be administered in a variety of dosage forms. Thus, they can be administered orally, for example as tablets, troches, lozenges, aqueous or oily suspensions, dispersible powders or granules. The inhibitors may also be administered parenterally, either subcutaneously, intravenously, intramuscularly, intrasternally, transdermally or by infusion techniques. The inhibitors may also be administered as suppositories. A physician will be able to determine the required route of administration for each particular patient.

The formulation of an inhibitor for use in preventing or treating a bacterial or fungal infection will depend upon factors such as the nature of the exact inhibitor, whether a pharmaceutical or veterinary use is intended, etc. An inhibitor may be

formulated for simultaneous, separate or sequential use.

An inhibitor is typically formulated for administration in the present invention with a pharmaceutically acceptable carrier or diluent. The pharmaceutical carrier or diluent may be, for example, an isotonic solution. For example, solid oral forms may contain, together with the active compound, diluents, e.g. lactose, dextrose, saccharose, cellulose, corn starch or potato starch; lubricants, e.g. silica, talc, stearic acid, magnesium or calcium stearate, and/or polyethylene glycols; binding agents; e.g. starches, gum arabic, gelatin, methylcellulose, carboxymethylcellulose or polyvinyl pyrrolidone; disaggregating agents, e.g. starch, alginic acid, alginates or sodium starch glycolate; effervescing mixtures; dyestuffs; sweeteners; wetting agents, such as lecithin, polysorbates, laurylsulphates; and, in general, non-toxic and pharmacologically inactive substances used in pharmaceutical formulations. Such pharmaceutical preparations may be manufactured in known manner, for example, by means of mixing, granulating, tabletting, sugar-coating, or film-coating processes.

Liquid dispersions for oral administration may be syrups, emulsions or suspensions. The syrups may contain as carriers, for example, saccharose or saccharose with glycerine and/or mannitol and/or sorbitol.

Suspensions and emulsions may contain as carrier, for example a natural gum, agar, sodium alginate, pectin, methylcellulose, carboxymethylcellulose, or polyvinyl alcohol. The suspensions or solutions for intramuscular injections may contain, together with the active compound, a pharmaceutically acceptable carrier, e.g. sterile water, olive oil, ethyl oleate, glycols, e.g. propylene glycol, and if desired, a suitable amount of lidocaine hydrochloride.

Solutions for intravenous administration or infusion may contain as carrier, for example, sterile water or preferably they may be in the form of sterile, aqueous, isotonic saline solutions.

A therapeutically effective amount of an inhibitor is administered to a patient. The dose of an inhibitor may be determined according to various parameters, especially according to the substance used; the age, weight and condition of the

patient to be treated; the route of administration; and the required regimen. Again, a physician will be able to determine the required route of administration and dosage for any particular patient. A typical daily dose is from about 0.1 to 50 mg per kg of body weight, according to the activity of the specific inhibitor, the age, weight and conditions of the subject to be treated, the type and severity of the degeneration and the frequency and route of administration. Preferably, daily dosage levels are from 5 mg to 2 g.

Conditional essential genes are good candidates for use in the preparation of live attenuated vaccines. The principle behind vaccination is to induce an immune response in the host thus providing protection against subsequent challenge with a pathogen. This may be achieved by inoculation with a live attenuated strain of the pathogen, i.e. a strain having reduced virulence such that it does not cause the disease caused by the virulent pathogen. Bacteria which carry mutations in conditional essential genes required for survival (i.e. growth/proliferation) in a host isolated according to the methods described above may be good candidates for use in live attenuated vaccines.

Mutations introduced into a bacterium for use in a vaccine generally knock-out the function of the conditional essential gene, for example a gene required for growth/proliferation in a host, completely. This may be achieved either by abolishing synthesis of any polypeptide at all from the gene or by making a mutation that results in synthesis of non-functional polypeptide. In order to abolish synthesis of polypeptide, either the entire gene or its 5'-end may be deleted. A deletion or insertion within the coding sequence of a gene may be used to create a gene that synthesises only non-functional polypeptide (e.g. polypeptide that contains only the N-terminal sequence of the wild-type protein).

The bacterium may have mutations in one or more, for example two, three or four conditional essential genes. The mutations are non-reverting mutations. These are mutations that show essentially no reversion back to the wild-type when the bacterium is used as a vaccine. Such mutations include insertions and deletions. Insertions and deletions are preferably large, typically at least 10 nucleotides in

length, for example from 10 to 600 nucleotides. Preferably, the whole coding sequence is deleted.

The bacterium used in the vaccine preferably contains only defined mutations, i.e. mutations which are characterised. It is clearly undesirable to use a bacterium which has uncharacterised mutations in its genome as a vaccine because there would be a risk that the uncharacterised mutations may confer properties on the bacterium that cause undesirable side-effects.

The attenuating mutations may be introduced by methods well known to those skilled in the art. Appropriate methods include cloning the DNA sequence of the wild-type gene into a vector, e.g. a plasmid, and inserting a selectable marker into the cloned DNA sequence or deleting a part of the DNA sequence, resulting in its inactivation. A deletion may be introduced by, for example, cutting the DNA sequence using restriction enzymes that cut at two points in or just outside the coding sequence and ligating together the two ends in the remaining sequence with an antibiotic resistance determinant. A plasmid carrying the inactivated DNA sequence can be transformed into the bacterium by known techniques such as electroporation or conjugation for example. It is then possible by suitable selection to identify a mutant wherein the inactivated DNA sequence has recombined into the chromosome of the bacterium and the wild-type DNA sequence has been rendered non-functional by homologous recombination.

The attenuated bacterium of the invention may be genetically engineered to express an antigen that is not expressed by the native bacterium (a "heterologous antigen"), so that the attenuated bacterium acts as a carrier of the heterologous antigen. The antigen may be from another organism, so that the vaccine provides protection against the other organism. A multivalent vaccine may be produced which not only provides immunity against the virulent parent of the attenuated bacterium but also provides immunity against the other organism. Furthermore, the attenuated bacterium may be engineered to express more than one heterologous antigen, in which case the heterologous antigens may be from the same or different organisms. The heterologous antigen may be a complete protein or a part of a protein containing

an epitope. The antigen may be from a virus, prokaryote or a eukaryote, for example another bacterium, a yeast, a fungus or a eukaryotic parasite. The antigen may be from an extracellular or intracellular protein. More especially, the antigenic sequence may be from *E.coli*, tetanus, hepatitis A, B or C virus, human rhinovirus such as type 2 or type 14, herpes simplex virus, poliovirus type 2 or 3, foot-and-mouth disease virus, influenza virus, coxsackie virus or Chlamydia trachomatis. Useful antigens include non-toxic components of E.coli heat labile toxin, *E.coli* K88 antigens, ETEC colonization factor antigens, P.69 protein from *B.pertussis* and tetanus toxin fragment C.

The DNA encoding the heterologous antigen is expressed from a promoter that is active in vivo. Two promoters that have been shown to work well in *Salmonella* are the *nir*B promoter and the *htr*A promoter. For expression of the ETEC colonization factor antigens, the wild-type promoters could be used. A DNA construct comprising the promoter operably linked to DNA encoding the heterologous antigen may be made and transformed into the attenuated bacterium using conventional techniques. Transformants containing the DNA construct may be selected, for example by screening for a selectable marker on the construct. Bacteria containing the construct may be grown in vitro before being formulated for administration to the host for vaccination purposes.

The vaccine may be formulated using known techniques for formulating attenuated bacterial vaccines. The vaccine is advantageously presented for oral administration, for example in a lyophilised encapsulated form. Such capsules may be provided with an enteric coating comprising, for example, Eudragate "S" (Trade Mark), Eudragate "L" (Trade Mark), cellulose acetate, cellulose phthalate or hydroxypropylmethyl cellulose. These capsules may be used as such, or alternatively, the lyophilised material may be reconstituted prior to administration, e.g. as a suspension. Reconstitution is advantageously effected in a buffer at a suitable pH to ensure the viability of the bacteria. In order to protect the attenuated bacteria and the vaccine from gastric acidity, a sodium bicarbonate preparation is advantageously administered before each administration of the vaccine. Alternatively, the vaccine

may be prepared for parenteral administration, intranasal administration or intramuscular administration.

The vaccine may be used in the vaccination of a mammalian host, particularly a human host but also an animal host. An infection caused by a microorganism, especially a pathogen, may therefore be prevented by administering an effective dose of a vaccine prepared according to the invention. The dosage employed will ultimately be at the discretion of the physician, but will be dependent on various factors including the size and weight of the host and the type of vaccine formulated. However, a dosage comprising the oral administration of from 10^7 to 10^{11} bacteria per dose may be convenient for a 70 kg adult human host.

Inhibitors of bacterial, fungal and pest essential or conditional essential genes may be administered to plants in order to prevent or treat bacterial, fungal or pest infections; the term pest includes any animal which attacks a plant. Thus inhibitors of the invention may be useful as pesticides. Inhibitors of plant essential or conditional essential genes may be administered to plants in order to reduce or stop plant growth, that is to act as a herbicide.

The inhibitors of the present invention are normally applied in the form of compositions together with one or more agriculturally acceptable carriers or diluents and can be applied to the crop area or plant to be treated, simultaneously or in succession with further compounds.

The inhibitors of the invention can be selective herbicides, bacteriocides, fungicides or pesticides or mixtures of several of these preparations, if desired together with further carriers, surfactants or application-promoting adjuvants customarily employed in the art of formulation. Suitable carriers and diluents correspond to substances ordinarily employed in formulation technology, e.g. natural or regenerated mineral substances, solvents, dispersants, wetting agents, tackifiers, binders or fertilizers.

A preferred method of applying active ingredients of the present invention or an agrochemical composition which contains at least one of the active ingredients is leaf application. The number of applications and the rate of application depend on

the intensity of infestation by the pathogen. However, the active ingredients can also penetrate the plant through the roots via the soil (systemic action) by impregnating the locus of the plant with a liquid composition, or by applying the compounds in solid form to the soil, e.g. in granular form (soil application). The active ingredients may also be applied to seeds (coating) by impregnating the seeds either with a liquid formulation containing active ingredients, or coating them with a solid formulation. In special cases, further types of application are also possible, for example, selective treatment of the plant stems or buds.

The active ingredients are used in unmodified form or, preferably, together with the adjuvants conventionally employed in the art of formulation, and are therefore formulated in known manner to emulsifiable concentrates, coatable pastes, directly sprayable or dilutable solutions, dilute emulsions, wettable powders, soluble powders, dusts, granulates, and also encapsulations, for example, in polymer substances. Like the nature of the compositions, the methods of application, such as spraying, atomizing, dusting, scattering or pouring, are chosen in accordance with the intended objectives and the prevailing circumstances. Advantageous rates of application are normally from 50g to 5kg of active ingredient (a.i.) per hectare ("ha", approximately 2.471 acres), preferably from 100g to 2kg a.i./ha, most preferably from 200g to 500g a.i./ha.

The formulations, compositions or preparations containing the active ingredients and, where appropriate, a solid or liquid adjuvant, are prepared in known manner, for example by homogeneously mixing and/or grindingactive ingredients with extenders, for example solvents, solid carriers and, where appropriate, surface-active compounds (surfactants).

Suitable solvents include aromatic hydrocarbons, preferably the fractions having 8 to 12 carbon atoms, for example, xylene mixtures or substituted naphthalenes, phthalates such as dibutyl phthalate or dioctyl phthalate, aliphatic hydrocarbons such as cyclohexane or paraffins, alcohols and glycols and their ethers and esters, such as ethanol, ethylene glycol, monomethyl or monoethyl ether, ketones such as cyclohexanone, strongly polar solvents such as N-methyl-2-pyrrolidone,

dimethyl sulfoxide or dimethyl formamide, as well as epoxidized vegetable oils such as epoxidized coconut oil or soybean oil; or water.

The solid carriers used e.g. for dusts and dispersible powders, are normally natural mineral fillers such as calcite, talcum, kaolin, montmorillonite or attapulgite. In order to improve the physical properties it is also possible to add highly dispersed silicic acid or highly dispersed absorbent polymers. Suitable granulated adsorptive carriers are porous types, for example pumice, broken brick, sepiolite or bentonite; and suitable nonsorbent carriers are materials such as calcite or sand. In addition, a great number of pregranulated materials of inorganic or organic nature can be used, e.g. especially dolomite or pulverized plant residues.

Depending on the nature of the active ingredient to be used in the formulation, suitable surface-active compounds are nonionic, cationic and/or anionic surfactants having good emulsifying, dispersing and wetting properties. The term "surfactants" will also be understood as comprising mixtures of surfactants. Suitable anionic surfactants can be both water-soluble soaps and water-soluble synthetic surface-active compounds.

Suitable soaps are the alkali metal salts, alkaline earth metal salts or unsubstituted or substituted ammonium salts of higher fatty acids (chains of 10 to 22 carbon atoms), for example the sodium or potassium salts of oleic or stearic acid, or of natural fatty acid mixtures which can be obtained for example from coconut oil or tallow oil. The fatty acid methyltaurin salts may also be used.

More frequently, however, so-called synthetic surfactants are used, especially fatty sulfonates, fatty sulfates, sulfonated benzimidazole derivatives or alkylarylsulfonates.

The fatty sulfonates or sulfates are usually in the form of alkali metal salts, alkaline earth metal salts or unsubstituted or substituted ammoniums salts and have a 8 to 22 carbon alkyl radical which also includes the alkyl moiety of alkyl radicals, for example, the sodium or calcium salt of lignonsulfonic acid, of dodecylsulfate or of a mixture of fatty alcohol sulfates obtained from natural fatty acids. These compounds also comprise the salts of sulfuric acid esters and sulfonic acids of fatty

alcohol/ethylene oxide adducts. The sulfonated benzimidazole derivatives preferably contain 2 sulfonic acid groups and one fatty acid radical containing 8 to 22 carbon atoms. Examples of alkylarylsulfonates are the sodium, calcium or triethanolamine salts of dodecylbenzenesulfonic acid, dibutylnaphthalenesulfonic acid, or of a naphthalenesulfonic acid/formaldehyde condensation product. Also suitable are corresponding phosphates, e.g. salts of the phosphoric acid ester of an adduct of p-nonylphenol with 4 to 14 moles of ethylene oxide.

Non-ionic surfactants are preferably polyglycol ether derivatives of aliphatic or cycloaliphatic alcohols, or saturated or unsaturated fatty acids and alkylphenols, said derivatives containing 3 to 30 glycol ether groups and 8 to 20 carbon atoms in the (aliphatic) hydrocarbon moiety and 6 to 18 carbon atoms in the alkyl moiety of the alkylphenols.

Further suitable non-ionic surfactants are the water-soluble adducts of polyethylene oxide with polypropylene glycol, ethylenediamine propylene glycol and alkylpolypropylene glycol containing 1 to 10 carbon atoms in the alkyl chain, which adducts contain 20 to 250 ethylene glycol ether groups and 10 to 100 propylene glycol ether groups. These compounds usually contain 1 to 5 ethylene glycol units per propylene glycol unit.

Representative examples of non-ionic surfactants are nonylphenolpolyethoxyethanols, castor oil polyglycol ethers, polypropylene/polyethylene oxide adducts, tributylphenoxypolyethoxyethanol, polyethylene glycol and octylphenoxyethoxyethanol. Fatty acid esters of polyoxyethylene sorbitan and polyoxyethylene sorbitan trioleate are also suitable non-ionic surfactants.

Cationic surfactants are preferably quaternary ammonium salts which have, as N-substituent, at least one C_8-C_{22} alkyl radical and, as further substituents, lower unsubstituted or halogenated alkyl, benzyl or lower hydroxyalkyl radicals. The salts are preferably in the form of halides, methylsulfates or ethylsulfates, e.g. stearyltrimethylammonium chloride or benzyldi(2-chloroethyl)ethylammonium bromide.

The surfactants customarily employed in the art of formulation are described, for example, in "McCutcheon's Detergents and Emulsifiers Annual", MC Publishing Corp. Ringwood, New Jersey, 1979, and Sisely and Wood, "Encyclopaedia of Surface Active Agents," Chemical Publishing Co., Inc. New York, 1980.

The agrochemical compositions usually contain from about 0.1 to about 99% preferably about 0.1 to about 95%, and most preferably from about 3 to about 90% of the active ingredient, from about 1 to about 99.9%, preferably from about 1 to 99%, and most preferably from about 5 to about 95% of a solid or liquid adjuvant, and from about 0 to about 25%, preferably about 0.1 to about 25%, and most preferably from about 0.1 to about 20% of a surfactant.

Whereas commercial products are preferably formulated as concentrates, the end user will normally employ dilute formulations.

The following Example illustrates the invention:

Example

Materials and Methods

Unless indicated otherwise, the methods used are standard biochemical techniques. Examples of suitable general methodology textbooks include Sambrook *et al.*, Molecular Cloning, a Laboratory Manual (1989) and Ausubel *et al.*, Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.

E. coli strains were grown on L-Agar (Sigma) and in L-Broth (Sigma). Kanamycin was purchased from Sigma. *S. typhimurium* SL1344 was grown in Tryptic Soy Broth (TSB, Oxoid).

Construction of the transposon

The EZ:Tn R6k ori Kan transposon (Epicentre) was used as a PCR template

with oligonucleotides:

97

(5'-CAG*CTGTCTCTTATACACATCT***CCCTATAGTGAGTCGTATT**ACCCATA ATACCCATAATAGCTGTTTGCCA**gtcgactctagagg-3'); and

98

(5'-CAG*CTGTCTCTTATACACATCT***CTTCTATAGTGTCACCTAAA**TAGGGAT AACAGGGTAAT**Gaattcgttaatacagatgt-3').

ME are *italicised*, the RNA polymerase binding sites are in bold, and the homing endonuclease sites are underlined. These oligonucleotides incorporate the T7 RNA polymerase site with the homing endonuclease site PI-*Psp*I, and the SP6 RNA polymerase site with the homing endonuclease site for I-*Sce*I, respectively. PCR was carried out according to protocol for Roche Expand Hi-fidelity Kit.

A first round of PCR, to introduce the homing endonuclease and RNA polymerase sites, was carried out with an initial denaturation step at 96°C for 3 min, then 5 cycles of 96°C 30s, 25°C 90s, 72°C 2min 30s, then a further 25 cycles at 96°C 30s, 50°C 90s, 72°C 2min 30s, final extension of 4 minutes and then cooled to 4°C. The PCR product was separated by gel electrophoresis, and the band (corresponding to the expected product at 2044bp) cut out and Gel extracted according to the Qiagen gel extraction protocol. The PCR products were re-ligated (according to Gibco protocol for T4 DNA ligase) overnight at 16°C. The ligation mixture was then purified using the Qiagen Gel Extraction protocol and the circularised transposon eluted in 50 µl water. 1 µl of the Transposon DNA was then electroporated into 40 µl TransforMax™ EC100D™ *pir*$^+$ Electrocompetent *E. coli* (Epicentre; using 0.1 cm cuvettes at 200Ω, 25 µF and 20KV/cm), outgrowth was in 1 ml SOC medium (Gibco life technologies) for 1 h at 37°C. The transformation was then plated out on L-Agar (Sigma) plates containing Kanamycin at 30 µg/ml and incubated overnight at 37°C. 12 colonies were then picked and grown up overnight in LB-Kan 30 µg/ml and 1.5 ml of culture was used for DNA-mini-preps according to the Qiagen protocol. The DNA sequence of the resulting constructs was then determined on a Beckman CEQ

DNA sequencer, (using manufacturers recommended conditions) with kan-2-FP-1 forward primer (Epicentre) as a DNA sequencing primer. The sequence of clone 11(GK11) was found to contain a single base pair change in one of the ME sequences. To correct this sequence change, an experiment using a second round of PCR was designed with an oligonucleotide that corrected the single base pair change. The clone GK11 (harbouring a single-point mutation in the ME) was re-PCR'd using un-phosphorylated ME primers (96°C for 3min, followed by 30 cycles of 96°C for 30s, 45°C for 90s, 72°C for 150s). The PCR products were then cloned into pBAD-TOPO (Invitrogen) and transformed into E. coli TOP10 chemically competent cells according to the Invitrogen protocol. The transformation mix was plated out on L-Agar containing kanamycin at 30 μg/ml and incubated overnight at 37°C. 6 colonies were then picked and grown in L-Broth containing kanamycin at 30 μg/ml overnight at 37°C. DNA from 1.5 mL samples of these cultures was then prepared using the miniprep method (Qiagen QIAprep spin miniprep kit) according to the Qiagen protocol. The DNA sequences of the inserts from each of these 6 minipreps was then determined with oligonucleotides kan-2-FP-1 and R6kan-2-RP-1 (Epicentre) as sequencing primers according to Beckman CEQ protocol. One of the isolates, clone 6A was correct when sequenced, named pBAD-GK6A, and was selected for further study.

Transposon evaluation

(i) T7 and SP6 RNA polymerase evaluation

In order to evaluate the novel transposon, 0.2 μg of pBAD-GK6A was used as template in an *in vitro* transcription (IVT) reactions using both SP6 (SP6 Megascript kit; Ambion) and T7 (T7 Megashortscipt kit, Ambion) RNA polymerases, according to Ambion protocols. RNA produced was purified using the Qiagen RNeasy mini Kit, and the amount of RNA produced measured at A_{260}. Transcription was observed both with the SP6 and T7 IVT reactions.

(ii) Restriction digest of transposon DNA with homing endonucleases

In order to determine whether or not the homing endonuclease sites were functional in the GK transposon, plasmid DNA (0.5 µg) was digested with I-*Sce*I and PI-*Psp*I. Separation of the resulting DNA products by agarose gel electrophoresis showed a single linearised band of the correct size.

(iii) Transposition of *Salmonella typhimurium*

pBAD-GK6A (0.1 µg) was electroporated into 40 µl *S. typhimurium* SL1344 electrocompetent cells (SL1344 was grown to an OD 0.5 in 100 ml tryptic-soy broth (TSB from Oxoid) at 37°C. Cells were centrifuged at 5000 x g for 10 min at 4°C, washed three times in 50 ml 10% glycerol before a final re-suspension in 1 ml of 10% glycerol) using a 0.2 cm cuvette, 200Ω, 25 µF and 12kV/cm, outgrowth was in 1 ml SOC (Gibco) at 37°C for 1 h, the transformation was then plated onto L-Agar plates containing Kanamycin at 50 µg/ml plates and grown overnight at 37°C. A colony was then picked and grown in 2.5 litres L-Broth containing Kanamycin at 50 µg/ml overnight at 37°C. A Qiagen Qiafilter Mega plasmid kit was used to purify pBAD-GK6A.

Generation of *S. typhimurium* SL1344 mutants with Transposon GK6a

The transposasome complex was generated from pBAD-GK6A as follows. pBAD-GK6a (100 µg) of was digested with *Xmn*I and *Nco*I (NEB) in NEB buffer 2 (with 1 µg BSA/ ml), overnight at 37°C. The entire digest was then run out on a 0.8% agarose gel, the ~2kb band, the GK transposon, was gel extracted using the Qiagen gel extraction kit, and eluted from 1 spin column in 50 µl TE pH8.5. The GK transposasome complex was generated according to Epicentre protocols and electroporated into electrocompetent *S. typhimurium* SL1344 cells as described above. Following outgrowth the cells were subsequently plated out on Tryptic Soy Agar plates (TSA, Oxoid) containing Kanamycin at 50 µg/ml and incubated overnight at 37°C. A total of 480 mutants were picked and grown overnight in 2 ml

TSB (Oxoid) containing Kanamycin at 50 μg/ml and glycerol stocks made (20% glycerol TSB).

Recovery of transposon for sequencing (use of R6k origin of replication)

Transposons and adjacent flanking DNA, corresponding to genes that have been disrupted by transposon insertion can be recovered from mutant chromosomal DNA samples using the use of R6k origin of replication. Digestion of chromosomal DNA purified from *S. typhimurium* SL1344 GK mutants with a restriction enzyme that does not cut in the transposon, followed by circularising the fragments, and transformation into a *pir*$^+$ strain of *E. coli*, results in the "rescuing" of this DNA.

Mutants 1-50 were grown up individually in 2 ml TSB containing Kanamycin at 50 μg/ml overnight at 37°C. Samples (1.5ml) was used to prepare chromosomal DNA using the Qiagen DNeasy tissue kit. A total of 5 μl (0.5 μg) of each of the fifty chromosomal preps was digested with *EcoRV* (NEB) in a final volume of 20 μl, overnight at 37°C. The *EcoRV* was heat inactivated at 80°C for 20 minutes. The 20 μl digest was then religated in 100 μl final volume using Gibco T4 DNA ligase, 48 hours at 4°C. Each religation was then individually cleaned up using a Qiagen gel extraction spin column and eluted in 50 μl water. Ligations (4 μl, 0.04 μg) were electroporated into electrocompetent *pir*$^+$ *E. coli* (EC100D, Epicentre) according to Epicentre protocols and plated on L-Agar containing Kanamycin at 30 μg/ml and incubated overnight at 37°C. Colonies were obtained from 46 on the electroporations, and were subsequently grown up in 5 ml L-Broth containing Kanamycin at 30 μg/ml 37°C overnight. Plasmid DNA (2 μg) from these clones was sequenced according to the Beckman CEQ protocol using oligonucleotide 108 as a sequencing primer (T7 end of the transposon).

Generation of RNA run-offs using iPCR and IVT for target hybridisation to microarrays

The generation of labelled target from GK mutants can be achieved by

inverse PCR (iPCR) amplification of each end of the transposon followed by IVT reactions using either SP6 or T7 RNA polymerises. A pool of 96 *S. typhimurium* SL1344 mutants was inoculated into L-Broth (10 ml) containing Kanamycin at 50 µg/ml and grown overnight at 37°C statically. Chromosomal DNA (20 µg) was prepared from 1.5ml of culture using the Qiagen DNeasy Kit, and 5 µl (0.5 µg) digested individually with the restriction enzymes *Hae*III, *Hha*I, *Hpych4*IV and *Rsa*I (NEB) in their respective NEB buffers in a final volume of 20 µl, overnight at 37°C. The enzymes were then heat denatured at 65°C for *Hha*I, *Hpych4* IV and *Rsa*I and 80°C for *Hae*III for 20 min. Each 20 µl digest was then self-ligated with T4 DNA ligase (Gibco) in a 100 µl reaction at 4°C for 48 h. Amplification of the DNA flanking each end of the transposon was achieved by iPCR. iPCR reactions to amplify the SP6 end of the transposon are performed with:

oligonucleotide 107 (5'-CTACCCTGTGGAACACCTACATCT-3');

and one of either

oligonucleotide 115 (5'-ATTACCTCTTTCTCCGCACCCGAC-3'; *Rsa*I or *Hpych4*IV) or

oligonucleotide116 (5'-CGACATAGATCCGGAACATAATGG-3'; *Hae*III or *Hha*I),

depending on the restriction enzyme used (in brackets) to cut the chromosomal DNA. iPCR reactions to amplify the T7 end of the transposon were performed with

oligonucleotide 108 (5'- ACCTACAACAAAGCTCTCATCAACC -3')

and one of either

oligonucleotide 117 (5'- ACAACCTATTAATTTCCCCTCGTC -3'; *Rsa*I, *Hae*III or *Hha*I) or

oligonucleotide118 (5'- ATGTTGGAATTTAATCGCGGCCTC -3'; *Hpych4*IV),

depending on the restriction enzyme used (in brackets) to cut the chromosomal DNA. iPCR reactions were then set up using Qiagen Taq polymerase according to the Qiagen protocol. 4 µl (0.02 µg) of ligation was used as template for each iPCR reaction. The reactions were initially denatured at 94?C 3' followed by 30 cycles of 94°C 30s, 65°C 90s, 72°C 90s followed by 7min extension at 72°C and then cooling

to 4°C. Each of the 8 iPCR's were purified using a Qiagen Gel extraction kit and the DNA eluted in 5 µl water. Each iPCR product was then re-digested with its respective restriction enzyme in a final volume of 50 µl (NEB) overnight 37°C. The digests were then cleaned using a Qiagen Gel extraction kit (following the manufacturers recommended procedure) and the DNA eluted in 50 µl EB pH8.5.

Each digested iPCR product (2 µl) was used as a template for both T7 and SP6 in vitro transcription reactions according to the Ambion protocol. The RNA was cleaned using the Qiagen Rneasy kit and eluted in 50 µl of RNase free water that was then placed in a UV transparent 96 well plate and the absorbance at 260 nm measured.

Ligation capture recovery of 'Gene Kelly'(GK) transposon ends

A significant advantage of the GK transposon is that it permits the recovery of DNA fragments adjacent to the site of transposon insertion by a method that does not employ PCR, which is ligation capture. Essentially, because of the rarity of the I-*Sce*I and PI-*Psp*I homing endonuclease sites the T7 and SP6 promoter sites that are linked to these sites can be enriched from a pool of DNA by ligation of a biotinylated linker to the cut site followed by purification using Streptavidin linked magnetic beads. A ligation capture experiment was performed on pBAD-GK6A. Plasmid DNA (1 µg) was digested with PI-*Psp*I overnight and the resulting linearised DNA purified using a Qiagen Gel extraction kit. This DNA (400 ng) was then digested overnight with *Hae*III and subsequently dephosphorylated using Calf Intestinal Alkaline Phosphatase (Roche). Dephosphorylated DNA was then ligated overnight onto a biotinylated linker, generated by annealing
oligonucleotide 113
(5'-biotin-GACGACCTCAGTTACGGTACGATCGGCCACGTAGCTTAT-3') and
oligonucleotide 114
(5'-phosphate-GCTACGTGGCCGATCGTACCGTAACTGAGGTCGTC-3').

The ligation was purified using a Qiagen Gel extraction. Biotinylated DNA

was then extracted from the ligation using Streptavidin-linked magnetic particles (150 µg; Promega) according to the manufacturers protocol, and the beads finally resuspended in 8 µl of 1x *Hae*III restriction buffer containing 10 units of *Hae*III (NEB). The digestion was then incubated at 37°C for 2 hours to remove the T7 RNA promoter from the linker. The beads were removed and an IVT reaction performed on the supernatant using an Ambion T7 Megashortscript kit, according to the manufacturers instructions. IVT products were purified using a Qiagen RNeasy kit, and the products eluted in 50 µl water and read at A_{260}. A value of $0.09A_{260}$ (1Aunit = 40 µg/ml) was achieved indicating that 180ng of RNA had been synthesised.

Results and Discussion

From the sequence data obtained we were able to identify the transposon insertion point in the published Salmonella genome LT2. Figure 4 shows a graph showing the random distribution of the sites of GK transposon distribution in the LT2 genome for the 46 sequenced mutants.

SEQUENCE LISTING

<110> ARROW THERAPEUTICS LIMITED

<120> TRANSPOSON

<130> P.84806 GCW

<160> 11

<170> PatentIn version 3.0

<210> 1
<211> 2044
<212> DNA
<213> Artificial transposon

<400> 1

ctgtctctta tacacatctc ttctatagtg tcacctaaat agggataaca gggtaatgaa 60
ttcgttaata cagatgtagg tgttccacag ggtagccagc agcatcctgc gatgcagatc 120
cggatgccat ttcattacct ctttctccgc acccgacata gatccgaaga tcagcagttc 180
aacctgttga tagtacgtac taagctctca tgtttcacgt actaagctct catgtttaac 240
gtactaagct ctcatgttta acgaactaaa ccctcatggc taacgtacta agctctcatg 300
gctaacgtac taagctctca tgtttcacgt actaagctct catgtttgaa caataaaatt 360
aatataaatc agcaacttaa atagcctcta aggttttaag ttttataaga aaaaaaagaa 420
tatataaggc ttttaaagct tttaaggttt aacggttgtg gacaacaagc cagggatctg 480
ccatttcatt acctctttct ccgcacccga catagatccg gaacataatg gtgcagggcg 540
ctgacttccg cgtttccaga ctttacgaaa cacggaaacc gaagaccatt catgttgttg 600
ctcaggtcgc agacgttttg cagcagcagt cgcttcacgt tcgctcgcgt atcggtgatt 660
cattctgcta accagtaagg caaccccgcc agcctagccg ggtcctcaac gacaggagca 720
cgatcatgcg cacccgtggc caggacccaa cgctgcccga gatgcgccgc gtgcggctgc 780
tggagatggc ggacgcgatg gatatgttct gccaagggtt ggtttgcgca ttcacagggt 840
gtctcaaaat ctctgatgtt acattgcaca agataaaaat atatcatcat gaacaataaa 900
actgtctgct tacataaaca gtaatacaag gggtgttatg agccatattc aacgggaaac 960
gtcttgctcg aggccgcgat taaattccaa catggatgct gatttatatg ggtataaatg 1020
ggctcgcgat aatgtcgggc aatcaggtgc gacaatctat cgattgtatg ggaagcccga 1080
tgcgccagag ttgtttctga aacatggcaa aggtagcgtt gccaatgatg ttacagatga 1140
gatggtcaga ctaaactggc tgacggaatt tatgcctctt ccgaccatca agcattttat 1200
ccgtactcct gatgatgcat ggttactcac cactgcgatc cccggaaaaa cagcattcca 1260
ggtattagaa gaatatcctg attcaggtga aaatattgtt gatgcgctgg cagtgttcct 1320
gcgccggttg cattcgattc ctgtttgtaa ttgtcctttt aacagcgatc gcgtatttcg 1380
tctcgctcag gcgcaatcac gaatgaataa cggtttggtt gatgcgagtg attttgatga 1440
cgagcgtaat ggctggcctg ttgaacaagt ctggaaagaa atgcataaac ttttgccatt 1500
ctcaccggat tcagtcgtca ctcatggtga tttctcactt gataaccttа ttttgacga 1560
ggggaaatta taggttgta ttgatgttgg acgagtcgga atcgcagacc gataccagga 1620
tcttgccatc ctatggaact gcctcggtga gttttctcct tcattacaga aacggctttt 1680
tcaaaaatat ggtattgata atcctgatat gaataaattg cagtttcatt tgatgctcga 1740
tgagtttttc taatcagaat tggttaattg gttgtaacac tggcagagca ttacgctgac 1800
ttgacgggac ggcggctttg ttgaataaat cgaacttttg ctgagttgaa ggatcagatc 1860
acgcatcttc ccgacaacgc agaccgttcc gtggcaaagc aaaagttcaa aatcaccaac 1920
tggtccacct acaacaaagc tctcatcaac cgtggcgggg atcctctaga gtcgactggc 1980
aaacagctat tatgggtatt atgggtaata cgactcacta tagggagatg tgtataagag 2040
acag 2044

<210> 2
<211> 85
<212> DNA
<213> Artificial oligonucleotide

<400> 2
cagctgtctc ttatacacat ctccctatag tgagtcgtat tacccataat acccataata 60
gctgtttgcc agtcgactct agagg 85

<210> 3
<211> 80
<212> DNA
<213> Artificial oligonucleotide

<400> 3
cagctgtctc ttatacacat ctcttctata gtgtcaccta aatagggata acagggtaat 60
gaattcgtta atacagatgt 80

<210> 4
<211> 24
<212> DNA
<213> Artificial oligonucleotide

<400> 4
ctaccctgtg gaacacctac atct 24

<210> 5
<211> 24
<212> DNA
<213> Artificial oligonucleotide

<400> 5
attacctctt tctccgcacc cgac 24

<210> 6
<211> 24
<212> DNA
<213> Artificial oligonucleotide

<400> 6
cgacatagat ccggaacata atgg 24

<210> 7
<211> 25
<212> DNA
<213> Artificial oligonucleotide

<400> 7
acctacaaca aagctctcat caacc 25

<210> 8
<211> 24
<212> DNA
<213> Artificial oligonucleotide

<400> 8
acaacctatt aatttcccct cgtc 24

<210> 9
<211> 24
<212> DNA
<213> Artificial oligonucleotide

<400> 9
atgttggaat ttaatcgcgg cctc 24

<210> 10
<211> 39
<212> DNA
<213> Artificial oligonucleotide

<400> 10
gacgacctca gttacggtac gatcggccac gtagcttat 39

<210> 11
<211> 35
<212> DNA
<213> Artificial oligonucleotide

<400> 11
gctacgtggc cgatcgtacc gtaactgagg tcgtc 35

CLAIMS

1 A transposon which comprises:

(i) an RNA polymerase recognition site;

(ii) a homing endonuclease recognition site; and

(iii) a bacterial origin of replication.

2. A transposon according to claim 1 which comprises two RNA polymerase recognition sites.

3. A transposon according to claim 2, wherein the two RNA polymerase recognition sites are different.

4. A transposon according to claim 3, wherein the two different RNA polymerase recognition sites are a T7 RNA polymerase recognition site and an SP6 RNA polymerase recognition site.

5. A transposon according to any one of the preceding claims which comprises two homing endonuclease recognition sites.

6. A transposon according to claim 5, wherein the two homing endonuclease recognition sites are different.

7. A transposon according to claim 6, wherein the two different homing endonuclease recognition sites are an I-*Sce*I recognition site and a PI-*Psp*I recognition site.

8. A transposon according to any one of the preceding claims, wherein the bacterial origin of replication is the R6k origin of replication.

9. A transposon according to any one of the preceding claims which is a modified *Tn5* transposon.

10. A transposon according to any one of the preceding claims, which comprises an antibiotic resistance gene.

11. Use of a transposon according to any one of the preceding claims in a method for the identification of an essential or a conditional essential gene.

12. A method for identifying an essential gene of an organism, which method comprises:

(i) providing a library of transposon insertion mutants of the said organism, wherein the transposon is a transposon according to any one of claims 1 to 10;

(ii) isolating chromosomal DNA from the library of (i);

(iii) digesting the chromosomal DNA with a restriction endonuclease that is capable of cutting 5' of the RNA polymerase recognition site(s) in the transposon and 3' of the RNA polymerase recognition site(s) in the chromosomal DNA flanking the transposon;

(iv) transcribing RNA from the digested DNA;

(v) hybridising the transcribed RNA with an oligonucleotide array; and

(vi) identifying a probe on the oligonucleotide array which corresponds to an essential gene of the organism.

13. A method according to claim 12, which comprises the additional steps:

(iii)' self-ligating the digested DNA; and

(iii)" amplifying the self-ligated DNA by inverse PCR (iPCR);

14. A method according to claim 13, wherein the amplified DNA produced in step (iii)" is digested with the same restriction endonuclease as used in step (iii)

before being transcribed in step (iv).

15. A method according to claim 13 or 14, wherein two oligonucleotides which bind divergently to recognition sites 5' to the RNA polymerase recognition site are used to carry out iPCR in step (iii)".

16. A method according to any one of claims 12 to 15, wherein a labelled ribonucleotide is used in transcribing RNA in step (iv).

17. A method according to any one of claims 12 to 16, wherein:

(a) the transposon is a transposon according to claim 3 or 4;

(b) step (iv) is carried out by transcribing one aliquot of the digested DNA with one RNA polymerase and transcribing a second aliquot of the digested DNA with a second different RNA polymerase; and

(c) step (v) is carried out by hybridising the two transcribed RNA pools with the same oligonucleotide array or separately with two copies of the same oligonucleotide array.

18. A method according to claim 17, wherein the two aliquots of digested RNA are each transcribed using a different labelled ribonucleotide.

19. A method according to any one of claims 12 to 18, wherein:

(a) aliquots of the chromosomal DNA are digested separately with different restriction endonucleases in step (iii);

(b) each of the restriction endonucleases are capable of cutting 5' to the RNA polymerase recognition site(s) in the transposons and 3' to the RNA polymerase recognition site(s) in the chromosomal DNA flanking the transposons; and

(c) each aliquot is subsequently being treated separately in steps (iv) to (vi).

20. A method according to claim 19, wherein two or three aliquots of the chromosomal DNA are digested separately with different restriction endonucleases.

21. A method for identifying an essential gene of an organism, which method comprises:

(i) providing a library of transposon insertion mutants of the said organism, wherein the transposon is a transposon according to any one of claims 1 to 10;

(ii) isolating chromosomal DNA from the library of (i);

(iii) digesting the chromosomal DNA with a homing endonuclease;

(iv) ligating the digested DNA with a biotinylated linker;

(v) recovering the digested DNA using streptavidin-coated particles;

(vi) transcribing RNA from the recovered digested DNA;

(vii) hybridising the transcribed RNA with an oligonucleotide array; and

(viii) identifying a probe on the oligonucleotide array which corresponds to an essential gene of the organism.

22. A method according to any one of claims 21, wherein a labelled ribonucleotide is used in transcribing RNA in step (iv).

23. A method according to claim 21 or 22, wherein:

(a) the transposon is a transposon according to claim 3 or 4;

(b) step (vi) is carried out by transcribing one aliquot of the digested DNA with one RNA polymerase and transcribing a second aliquot of the digested DNA with a second different RNA polymerase; and

(c) step (vii) is carried out by hybridising the two transcribed RNA pools with the same oligonucleotide array or separately with two copies of the same oligonucleotide array.

24. A method according to claim 23, wherein the two aliquots of digested RNA are each transcribed using a different labelled ribonucleotide.

25. A method for identifying a conditional essential gene of an organism, which method comprises:

(a) providing a first sample of a library of transposon insertion mutants of the said organism (input library);

(b) providing a second sample of the library and subjecting that sample to a conditional restraint;

(c) collecting the mutants that survive the conditional restraint in step (ii) to give a second library (output library);

(d) carrying out a method according to steps (ii) to (iv) of any one of claims 12 to 20 or according to steps (ii) to (vi) of any one of claims 21 to 23 on the input library from step (a) and on the output library from step (c);

(e) hybridising the transcribed RNA derived from the input library and from the output library to the same oligonucleotide array or different separately to copies of the same oligonucleotide array; and

(f) identifying a probe on the oligonucleotide array(s) which corresponds to a conditional essential gene of the organism.

26. A method according to claim 25, wherein the organism is a bacterium and the conditional restraint is growth of that bacterium in its host.

27. A method according to any one of the preceding claims, wherein the oligonucleotide array comprises probes which are from 9 to 150 bp in length.

28. A method according to any one of the preceding claims, wherein the oligonucleotide array comprises 1 probe for every 60 to 250 bp of the locus or loci represented on the array.

29. Use of an essential or conditional essential gene identified by a method according to any one of claims 12 to 28, or a polypeptide encoded by a said gene, in a method for identifying an inhibitor of transcription and/or translation of that gene and/or activity of a polypeptide encoded that gene.

30. A method for identifying an inhibitor of transcription and/or translation of an essential or conditional essential gene and/or an inhibitor of activity of a polypeptide encoded by a said gene, which method comprises:

(a) identifying an essential or conditional essential gene by a method according to any one of claims 12 to 28; and

(b) determining whether a test substance can inhibit transcription and/or translation of a gene identified in step (a) and/or activity of a polypeptide encoded by a said identified gene, thereby to identify a said inhibitor.

31. An inhibitor identified by a method according to claim 30.

32. An inhibitor according to claim 31 for use in a method of treatment of a bacterial, fungal or eukaryotic parasite infection, wherein the essential or conditional essential gene in claim 30 is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene.

33. Use of an inhibitor according to claim 31 wherein the essential or conditional essential gene in claim 30 is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene, in the manufacture of a medicament for use in the treatment of a bacterial, fungal or eukaryotic parasite infection.

34. A pharmaceutical composition comprising an inhibitor according to claim 31 wherein the essential or conditional essential gene in claim 30 is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene and a pharmaceutically acceptable carrier or diluent.

35. A method of treating a host suffering from a bacterial, fungal or eukaryotic parasite infection, which method comprises the step of administering to the host a therapeutically effective amount of an inhibitor according to claim 31 wherein the essential or conditional essential gene in claim 30 is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene.

36. A method for the preparation of a pharmaceutical composition, which method comprises:

(a) identifying an inhibitor of transcription and/or translation of an essential or conditional essential gene of an organism and/or an inhibitor of activity of a polypeptide encoded by a said gene by a method according to claim 30 wherein the essential or conditional essential gene is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene; and

(b) formulating the inhibitor thus identified with a pharmaceutically acceptable carrier or diluent.

37. A method for treating a host suffering from a bacterial, fungal or eukaryotic parasite infection, which method comprises:

(a) identifying an inhibitor of transcription and/or translation of an essential or conditional essential gene of an organism and/or an inhibitor of activity of a polypeptide encoded by a said gene by a method according to claim 30 wherein the essential or conditional essential gene is a bacterial, fungal or eukaryotic parasite essential or conditional essential gene;

(b) formulating the inhibitor thus identified with a pharmaceutically acceptable carrier or diluent; and

(c) administering to the host a therapeutically effective amount of an inhibitor thus formulated.

38. An inhibitor according to claim 31, wherein the essential or conditional essential gene is a plant bacterial, plant fungal or plant pest essential or conditional

essential gene.

39. Use of an inhibitor according to 38 as a plant bactericide, fungicide or pesticide.

40. An inhibitor according to claim 31, wherein essential or conditional essential gene is a plant essential or conditional essential gene.

41. Use of an inhibitor according to claim 40 as a herbicide.

42. A bacterium attenuated by a non-reverting mutation in one or more genes identified by a method as defined in claim 30.

43. A vaccine comprising a bacterium according to claim 42 and a pharmaceutically acceptable carrier or diluent.

44. A bacterium according to claim 42 for use in a method of vaccinating a human or animal.

45. Use of a bacterium according to claim 42 for the manufacture of a medicament for vaccinating a human or animal.

46. A method for raising an immune response in a mammalian host, which method comprises the step of administering to the host a bacterium according to claim 42.

47. A method for preparing an attenuated bacterium, which method comprises:

(a) identifying a conditional essential gene in a bacterium by a method according to claim 30; and

(b) introducing a non-reverting mutation into a thus-identified conditional

essential gene of the bacterium, thereby to attenuate the bacterium.

48. A method for the preparation of a vaccine, which method comprises:

(a) identifying a conditional essential gene in a bacterium by a method according to claim 30;

(b) introducing a non-reverting mutation into a thus-identified conditional essential gene of the bacterium, thereby to attenuate the bacterium; and

(c) formulating the attenuated bacterium with a pharmaceutically acceptable carrier or diluent.

49. A method for raising an immune response in a mammalian host, which method comprises:

(a) identifying a conditional essential gene in a bacterium by a method according to claim 30;

(b) introducing a non-reverting mutation into a thus-identified conditional essential gene of the bacterium, thereby to attenuate the bacterium;

(c) formulating the attenuated bacterium with a pharmaceutically acceptable carrier or diluent; and

(d) administering to the host the attenuated bacterium thus formulated.

ABSTRACT

TRANSPOSON

A transposon comprises: (i) an RNA polymerase recognition site; (ii) a homing endonuclease recognition site; and (iii) a bacterial origin of replication. The transposon may be used in methods for the identification of an essential gene or a conditional essential gene of an organism. Genes identified in such methods are useful as substrates for use in screens for antibacterials, antiparasitics, fungicides, pesticides and herbicides.

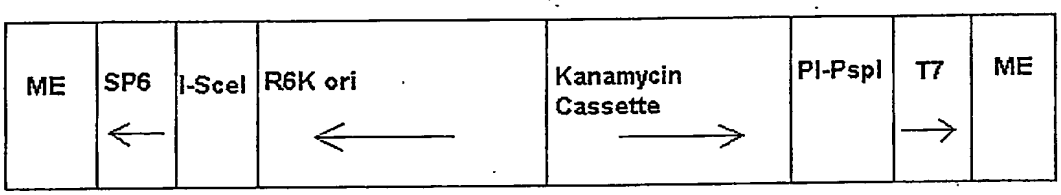


Fig. 1

*CTGTCTCTTATACACATCT*CTTCTATAGTGTCACCTAAATAGGGATAACAGGG
TAATGAATTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATC
CTGCGATGCAGATCCGGATGCCATTTCATTACCTCTTTCTCCGCACCCGACATA
GATCCGAAGATCAGCAGTTCAACCTGTTGATAGTACGTACTAAGCTCTCATGT
TTCACGTACTAAGCTCTCATGTTTAACGTACTAAGCTCTCATGTTTAACGAACT
AAACCCTCATGGCTAACGTACTAAGCTCTCATGGCTAACGTACTAAGCTCTCAT
GTTTCACGTACTAAGCTCTCATGTTTGAACAATAAAATTAATATAAATCAGCAA
CTTAAATAGCCTCTAAGGTTTTAAGTTTTATAAGAAAAAAAAAGAATATATAAG
GCTTTTAAAGCTTTTAAGGTTTAACGGTTGTGGACAACAAGCCAGGGATCTGC
CATTTCATTACCTCTTTCTCCGCACCCGACATAGATCCGGAACATAATGGTGCA
GGGCGCTGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCA
TTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTT
CGCTCGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCC
TAGCCGGGTCCTCAACGACAGGAGCACGATCATGCGCACCCGTGGCCAGGAC
CCAACGCTGCCCGAGATGCGCCGCGTGCGGCTGCTGGAGATGGCGGACGCG
ATGGATATGTTCTGCCAAGGGTTGGTTTGCGCATTCACAGGGTGTCTCAAAAT
CTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTG
TCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGA
AACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATG
GGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCG
ATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGT
AGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGG
AATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCAT
GGTTACTCACCACTGCGATCCCCGGAAAAACAGCATTCCAGGTATTAGAAGAA
TATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCG
GTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCG
TCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATT
TTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCA
TAAACTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACT
TGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGAC
GAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTC
GGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGAT
AATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA
TCAGAATTGGTTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACG
GGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGA
TCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAA
ATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCGGGGATCC
TCTAGAGTCGACTGGCAAACAGCTATTATGGGTATTATGGGTAATACGACTCA
CTATAGGG*AGATGTGTATAAGAGACAG*

Fig. 2

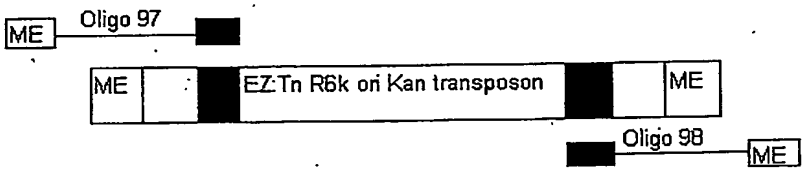


Fig. 3

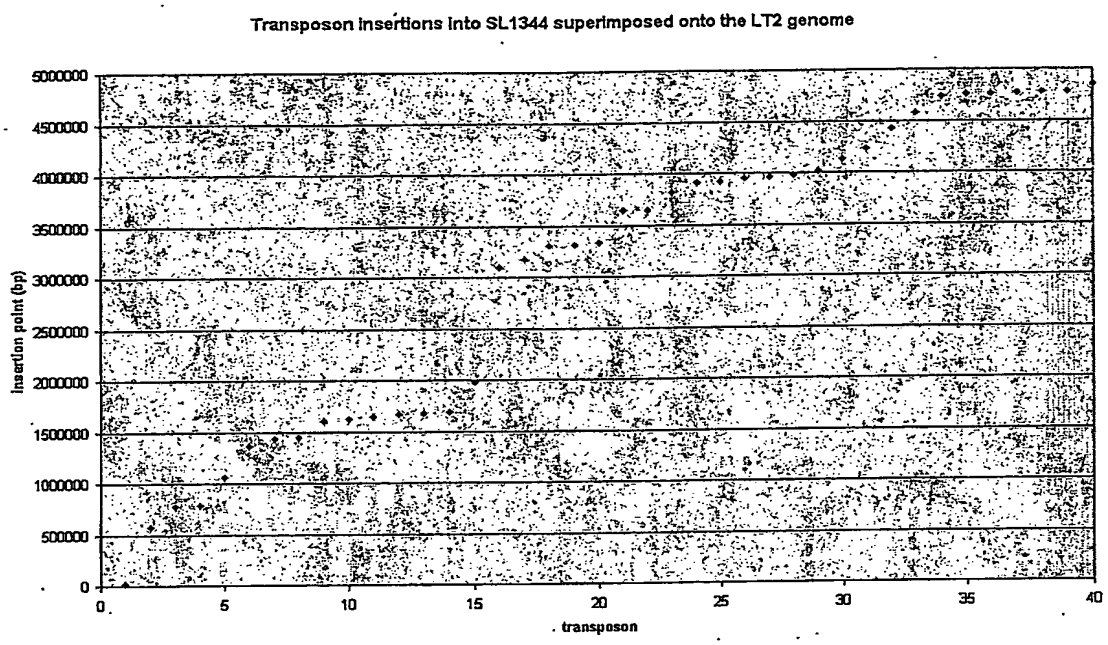

Transposon insertions into SL1344 superimposed onto the LT2 genome
Insertion point (bp)
5000000
4500000
4000000
3500000
3000000
2500000
2000000
1500000
1000000
500000
0
0
5
10
15
20
25
30
35
40
transposon

Fig. 4

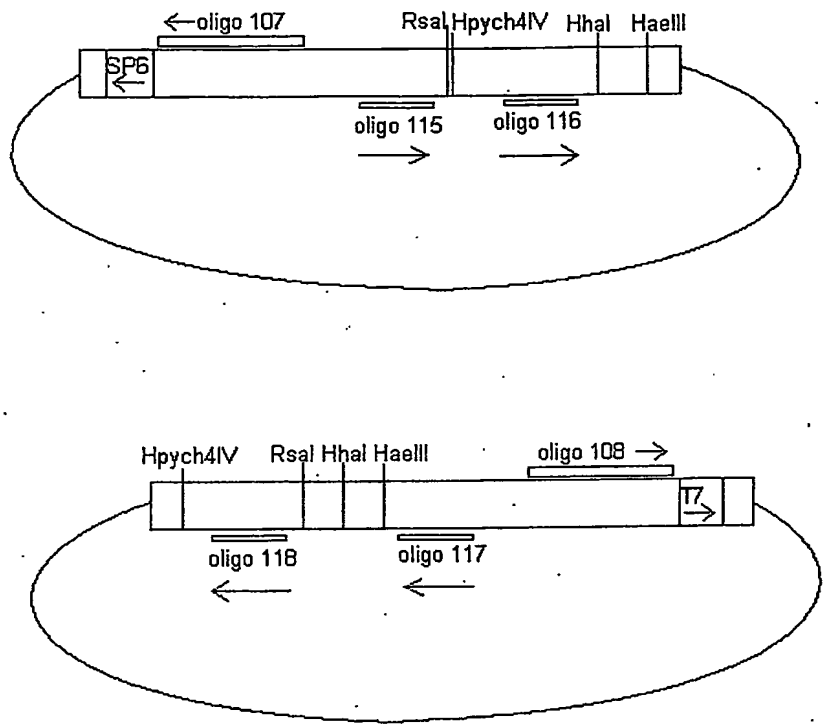

←oligo 107
RsaI
Hpych4IV
HhaI
HaeIII
SP6
oligo 115
oligo 116
Hpych4IV
RsaI
HhaI
HaeIII
oligo 108 →
T7
oligo 118
oligo 117

Fig. 5

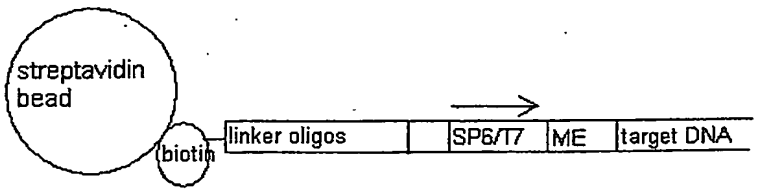

streptavidin
bead
biotin
linker oligos
SP6/T7
ME
target DNA

Fig. 6

PCT GB 2003 000918